GenCore version 4.5

OM protein - protein search, using sw model

Run on: June 4, 2002, 10:30:44 ; Search time 13.4 Seconds
(without alignments)
1112.465 Million cell updates/sec

Title: US-09-402-845-2
Perfect score: 2060
Sequence: 1 MGAVPLPLSPTASLSLGFLL.....DWATECMATSSHQGTVGALG 385

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID		Description
1	1818	88.3	381	1	PPAP_RAT	P20646	rattus norv
2	1661.5	80.7	386	1	PPAP_HUMAN	P15309	homo sapien
3	1022.5	49.6	423	1	PPAL_HUMAN	P11117	homo sapien
4	1020.5	49.5	421	1	PPAL_MOUSE	P24638	mus musculu
5	1019.5	49.5	423	1	PPAL_RAT	P20611	rattus norv
6	499.5	24.2	394	1	PPAY_CAEEL	Q10944	caenorhabdi
7	310	15.0	755	1	PPAX_CAEEL	Q09549	caenorhabdi
8	274	13.3	413	1	PPAW_CAEEL	Q09451	caenorhabdi
9	247.5	12.0	471	1	PPAV_CAEEL	Q09448	caenorhabdi
10	121	5.9	413	1	AGP_ECOLI	P19926	escherichia
11	115.5	5.6	737	1	VE02_VACCC	P21080	vaccinia vi
12	110.5	5.4	737	1	VE02_VACCV	P21604	vaccinia vi
13	110	5.3	599	1	RRPO_PEAMV	P29154	pea enation
14	109.5	5.3	737	1	VE02_VARV	P33862	variola vir
15	106	5.1	463	1	PPA2_SCHPO	Q01682	schizosacch
16	105	5.1	417	1	AGP_PRORE	Q52309	providencia
17	104.5	5.1	432	1	PPA_ECOLI	P07102	escherichia
18	104	5.0	463	1	PHYB_EMENI	O00093	emericella
19	104	5.0	489	1	PT94_YEAST	P07390	saccharomyc
20	102.5	5.0	479	1	PHYB_ASPAW	P34755	aspergillus
21	102	5.0	413	1	AGP_SALTY	O33921	salmonella
22	101	4.9	801	1	SUS2_DAUCA	O49845	daucus caro
23	99.5	4.8	937	1	MSH2_ARATH	O24617	arabidopsis
24	97.5	4.7	320	1	HLP1_HUMAN	P54257	homo sapien
25	97.5	4.7	453	1	PPA1_SCHPO	P08091	schizosacch
26	97.5	4.7	467	1	PHYA_ASPNG	P34752	aspergillus
27	97	4.7	290	1	Y290_LAMBD	P03766	bacteriopha
28	97	4.7	805	1	SUSY_PHAAU	Q01390	phaseolus a
29	97	4.7	805	1	SUSY_SOYBN	P13708	glycine max
30	96.5	4.7	503	1	CP39_RAT	P51538	rattus norv
31	96	4.7	653	1	MALQ_PYRKO	O32450	pyrococcus
32	95	4.6	820	1	SUS2_TULGE	Q41607	tulipa gesn
33	94	4.6	468	1	PPA1_PICPA	P52291	pichia past
34	94	4.6	728	1	Y376_METJA	Q57821	methanococc
35	94	4.6	806	1	SUSY_VICFA	P31926	vicia faba
36	93.5	4.5	1103	1	CHS6_USTMA	O13395	ustilago ma
37	93.5	4.5	1183	1	DRPL_RAT	P54258	rattus norv
38	93	4.5	465	1	TPSN_MOUSE	Q9r233	mus musculu
39	93	4.5	805	1	SUSY_MEDSA	O65026	medicago sa
40	92	4.5	655	1	AMYA_PYRAB	Q9v298	pyrococcus
41	91.5	4.4	452	1	HOS2_YEAST	P53096	saccharomyc
42	91.5	4.4	467	1	PHYA_ASPAW	P34753	aspergillus
43	91.5	4.4	1185	1	DRPL_HUMAN	P54259	homo sapien
44	91.5	4.4	2717	1	ZEP1_HUMAN	P15822	homo sapien
45	91	4.4	536	1	DIT1_YEAST	P21623	saccharomyc

ALIGNMENTS

RESULT 1
PPAP_RAT
ID PPAP_RAT STANDARD; PRT; 381 AA.
AC P20646;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE Prostatic acid phosphatase precursor (EC 3.1.3.2).
GN ACPP.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90323620; PubMed=2373368;
RA Roiko K., Jaenne O.A., Vihko P.;
RT "Primary structure of rat secretory acid phosphatase and comparison
RT to other acid phosphatases.";
RL Gene 89:223-229(1990).
RN [2]
RP X-RAY CRYSTALLOGRAPHY (3.0 ANGSTROMS).
RX MEDLINE=93327749; PubMed=8334986;
RA Schneider G., Lindqvist Y., Vihko P.;
RT "Three-dimensional structure of rat acid phosphatase.";
RL EMBO J. 12:2609-2615(1993).
RN [3]
RP X-RAY CRYSTALLOGRAPHY (3.0 ANGSTROMS).
RX MEDLINE=94012606; PubMed=8407898;
RA Lindqvist Y., Schneider G., Vihko P.;
RT "Three-dimensional structure of rat acid phosphatase in complex with
RT L(+)-tartrate.";
RL J. Biol. Chem. 268:20744-20746(1993).
CC -!- CATALYTIC ACTIVITY: An orthophosphoric monoester + H(2)O = an
CC alcohol + phosphate.
CC -!- SUBUNIT: HOMODIMER.
CC -!- SIMILARITY: BELONGS TO THE HISTIDINE ACID PHOSPHATASE FAMILY.
CC ---
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC ---
DR EMBL; M32397; AAA41806.1; -.
DR PIR; JH0152; JH0152.
DR PDB; 1RPA; 31-MAY-94.
DR PDB; 1RPT; 31-MAY-94.
DR InterPro; IPR000560; His_acid_phosphtse.
DR Pfam; PF00328; acid_phosphat; 1.
DR PROSITE; PS00778; HIS_ACID_PHOSPHAT_2; FALSE_NEG.
DR PROSITE; PS00616; HIS_ACID_PHOSPHAT_1; 1.
KW Hydrolase; Glycoprotein; Signal; 3D-structure.
FT SIGNAL 1 31

```
FT   CHAIN        32    381       PROSTATIC ACID PHOSPHATASE.
FT   DISULFID    160    371
FT   DISULFID    346    350
FT   ACT_SITE     43     43
FT   ACT_SITE     85     85       BY SIMILARITY.
FT   CARBOHYD     93     93       N-LINKED (GLCNAC. . .).
FT   CARBOHYD    219    219       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    332    332       N-LINKED (GLCNAC. . .).
SQ   SEQUENCE   381 AA;  43850 MW;  5EEBFF67B062FF76 CRC64;


  Query Match             88.3%;  Score 1818;  DB 1;  Length 381;
  Best Local Similarity   88.5%;  Pred. No. 3.7e-137;
  Matches  337;  Conservative   24;  Mismatches   20;  Indels    0;  Gaps    0;

Qy       1 MGAVPLPLSPTASLSLGFLLLLSLCLDPGQAKELKFVTLVFRHGDRGPIETFPTDPITES 60
           | ||||  |  ||||:||||||||| |||||||||||||||||||||||||||| ||| ||
Db       1 MRAVPLHLVGTASLTLGFLLLLSLRLDPGQAKELKFVTLVFRHGDRGPIETFPNDPIKES 60

Qy      61 SWPQGFGQLTQWGMEQHYELGSYIRKRYGRFLNDTYKHDQIYIRSTDVDRTLMSAMTNLA 120
           ||||||||||:||| ||||||||||:|||||||::|||||:|||||||||||||||||||
Db      61 SWPQGFGQLTKWGMGQHYELGSYIRRRYGRFLNNSYKHDQVYIRSTDVDRTLMSAMTNLA 120

Qy     121 ALFPPEGISIWNPRLLWQPIPVHTVSLSEDRLLYLPFRDCPRFEELKSETLESEEFLKRL 180
           ||||||||||||||||||||||||||||||||||||||||||:|||||||:||||||||||
Db     121 ALFPPEGISIWNPRLLWQPIPVHTVSLSEDRLLYLPFRDCPRFQELKSETLKSEEFLKRL 180

Qy     181 HPYKSFLDTLSSLSGFDDQDLFGIWSKVYDPLFCESVHNFTLPSWATEDAMIKLKELSEL 240
           |||||:||| |||||:||||| |||::||||:||||||||  :|||||||| ||||||||
Db     181 QPYKSFIDTLPSLSGFEDQDLFEIWSRLYDPLYCESVHNFTFRTWATEDAMTKLKELSEL 240

Qy     241 SLLSLYGIHKQKEKSRLQGGVLVNEILKNMKLATQPQKYKKLVMYSAHDTTVSGLQMALD 300
           ||||||||||||||||||||||||||||||||| :||:||||:||||||||||:
Db     241 SLLSLYGIHKQKEKSRLQGGVLVNEILKNMKLATQPQKARKLIMYSAYDTTVSGLQMALE 300

Qy     301 VYNGVLPPYASCHMMELYHDKGGHFVEMYYRNETQNEPYPLTLPGCTHSCPLEKFAELLD 360
           :|||:||||||||:|||| | || ||||||||||||||||||||||||||||||||||||
Db     301 LYNGLLPPYASCHIMELYQDNGGTFVEMYYRNETQNEPYPLTLPGCTHSCPLEKFAELLD 360

Qy     361 PVIPQDWATECMATSSHQGTV 381
           |||||||||||| ||:|| ::
Db     361 PVIPQDWATECMGTSNHQASL 381


RESULT   2
PPAP_HUMAN
ID   PPAP_HUMAN     STANDARD;      PRT;   386 AA.
AC   P15309;
DT   01-APR-1990 (Rel. 14, Created)
DT   01-MAR-1992 (Rel. 21, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Prostatic acid phosphatase precursor (EC 3.1.3.2).
GN   ACPP.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=92272747; PubMed=1375464;
RA   Sharief F.S., Li S.S.-L.;
RT   "Structure of human prostatic acid phosphatase gene.";
RL   Biochem. Biophys. Res. Commun. 184:1468-1476(1992).
RN   [2]
RP   SEQUENCE FROM N.A., PARTIAL SEQUENCE, DISULFIDE BONDS, AND ACTIVE
RP   SITE.
RX   MEDLINE=91115848; PubMed=1989985;
RA   van Etten R.L., Davidson R., Stevis P.E., Macarthur H., Moore D.L.;
RT   "Covalent structure, disulfide bonding, and identification of
RT   reactive surface and active site residues of human prostatic acid
RT   phosphatase.";
RL   J. Biol. Chem. 266:2313-2319(1991).
RP   SEQUENCE FROM N.A.
RX   MEDLINE=89228054; PubMed=2712834;
RA   Sharief F.S., Lee H., Leuderman M.M., Lundwall A., Deaven L.L.,
RA   Lee C.-L., Li S.S.-L.;
RT   "Human prostatic acid phosphatase: cDNA cloning, gene mapping and
RT   protein sequence homology with lysosomal acid phosphatase.";
RL   Biochem. Biophys. Res. Commun. 160:79-86(1989).
RN   [4]
RP   SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC   TISSUE=Prostate;
RX   MEDLINE=88312981; PubMed=2842184;
RA   Vihko P., Virkkunen P., Henttu P., Roiko K., Solin T., Huhtala M.L.;
RT   "Molecular cloning and sequence analysis of cDNA encoding human
RT   prostatic acid phosphatase.";
RL   FEBS Lett. 236:275-281(1988).
RN   [5]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Prostate;
RX   MEDLINE=90370491; PubMed=2395659;
RA   Tailor P.G., Govindan M.V., Patel P.C.;
RT   "Nucleotide sequence of human prostatic acid phosphatase determined
RT   from a full-length cDNA clone.";
RL   Nucleic Acids Res. 18:4928-4928(1990).
RN   [6]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=95038536; PubMed=7951074;
RA   Sharief F.S., Li S.S.-L.;
RT   "Nucleotide sequence of human prostatic acid phosphatase ACPP gene,
RT   including seven Alu repeats.";
RL   Biochem. Mol. Biol. Int. 33:561-565(1994).
RN   [7]
RP   X-RAY CRYSTALLOGRAPHY (2.9 ANGSTROMS).
RX   MEDLINE=99023966; PubMed=9804805;
RA   Lacount M.W., Handy G., Lebioda L.;
RT   "Structural origins of L(+)-tartrate inhibition of human prostatic
RT   acid phosphatase.";
RL   J. Biol. Chem. 273:30406-30409(1998).
CC   -!- CATALYTIC ACTIVITY: An orthophosphoric monoester + H(2)O = an
CC       alcohol + phosphate.
CC   -!- SUBUNIT: HOMODIMER.
CC   -!- SIMILARITY: BELONGS TO THE HISTIDINE ACID PHOSPHATASE FAMILY.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; M97589; AAA60021.1; -.
DR   EMBL; M97580; AAA60021.1; JOINED.
DR   EMBL; M97581; AAA60021.1; JOINED.
DR   EMBL; M97582; AAA60021.1; JOINED.
DR   EMBL; M97583; AAA60021.1; JOINED.
DR   EMBL; M97584; AAA60021.1; JOINED.
DR   EMBL; M97585; AAA60021.1; JOINED.
DR   EMBL; M97586; AAA60021.1; JOINED.
DR   EMBL; M97587; AAA60021.1; JOINED.
DR   EMBL; M97588; AAA60021.1; JOINED.
DR   EMBL; M34840; AAA69694.1; -.
DR   EMBL; M24902; AAA60022.1; -.
DR   EMBL; X52174; CAA36422.1; -.
DR   EMBL; X53605; CAA37673.1; -.
DR   EMBL; U07097; AAB60640.1; -.
DR   EMBL; U07083; AAB60640.1; JOINED.
DR   EMBL; U07085; AAB60640.1; JOINED.
DR   EMBL; U07086; AAB60640.1; JOINED.
DR   EMBL; U07088; AAB60640.1; JOINED.
DR   EMBL; U07091; AAB60640.1; JOINED.
DR   EMBL; U07092; AAB60640.1; JOINED.
DR   EMBL; U07093; AAB60640.1; JOINED.
DR   EMBL; U07095; AAB60640.1; JOINED.
```

```
Qy    363 IPQDWATECMAT 374
          |: :   | |
Db    408 SPEKYHALCSQT 419
```

RESULT 12
US-10-121-060-330
; Sequence 330, Application US/10121060
; GENERAL INFORMATION:
; APPLICANT: Baker,Kevin P.
; APPLICANT: Beresini,Maureen
; APPLICANT: DeForge,Laura
; APPLICANT: Desnoyers,Luc
; APPLICANT: Filvaroff,Ellen
; APPLICANT: Gao,Wei-Qiang
; APPLICANT: Gerritsen,Mary E.
; APPLICANT: Goddard,Audrey
; APPLICANT: Godowski,Paul J.
; APPLICANT: Gurney,Austin L.
; APPLICANT: Sherwood,Steven
; APPLICANT: Smith,Victoria
; APPLICANT: Stewart,Timothy A.
; APPLICANT: Tumas,Daniel
; APPLICANT: Watanabe,Colin K
; APPLICANT: Wood,William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3330R1C21
; CURRENT APPLICATION NUMBER: US/10/121,060
; CURRENT FILING DATE: 2002-04-12
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 330
; LENGTH: 428
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-121-060-330

Query Match 15.3%; Score 315.5; DB 6; Length 428;
Best Local Similarity 26.9%; Pred. No. 8.1e-20;
Matches 116; Conservative 62; Mismatches 155; Indels 99; Gaps 20;

```
Qy     16 LGFLLLLSLCLD------------PGQAK------ELKFVTLVFRHGDRGPIETFPTD-- 55
          :| |  |: ||              ||        :|| | :||||| | |::  | :
Db     14 VGVLTSLAYCLHQRRVALAELQEADGQCPVDRSLLKLKMVQVVFRHGARSPLKPLPLEEQ 73

Qy     56 -----------PITESSW--------PQGF-----------------GQLTQWGMEQHY 78
                      | |:  :        |: :                 ||||: ||:| :
Db     74 VEWNPQLLEVPPQTQFDYTVTNLAGGPKPYSPYDSQYHETTLKGGMFAGQLTKVGMQQMF 133

Qy     79 ELGSYIRKRYGR---FLNDTYKHDQIYIRSTDVDRTLMSAMTNLAALF--PPEGISIWNP 133
           ||  :|| |     ||: |:   :::||||:: | | |    || ||    ||
Db    134 ALGERLRKNYVEDIPFLSPTFNPQEVFIRSTNIFRNLESTRCLLAGLFQCQKEG------ 187

Qy    134 RLLWQPIPVHTVSLSEDRLLYLPFRDCPRFEELKSETLESEEFLKRLHPYKSFLDTLSSL 193
               || :||   ::  :|| :: |     |: |    :                | :
Db    188 -----PIIIHT-DEADSEVLYPNYQSC---WSLRQRTRGRRQTASLQPGISEDLKKVKDR 238

Qy    194 SGFDDQDLFGIWSKVYDPLFCESVHNFTLPSWATEDAMIKLKELSELSLLSLYGIHKQ-K 252
           | |  |     :  : | :  |  ||  |||         :: |   :   ||| : |: :
Db    239 MGIDSSDKVDFFI-LLDNVAAEQAHN--LPSCPMLKRFARMIEQRAVD-TSLYILPKEDR 294

Qy    253 EKSRLQGG----VLVNEILKNMKLATQPQKYKKLVMYSAHDTTVSGLQMALDVYNGVLPP 308
          |  ::  |    :| : :|| |  || | | :|| :|:||| |   | | | :::   ||
Db    295 ESLQMAVGPFLHILESNLLKAMDSATAPDKIRKLYLYAAHDVTFIPLLMTLGIFDHKWPP 354

Qy    309 YASCHMMELYH--DKGGHFVEMYYRNETQNEPYPLTLP-GCTHS-CPLEKFAELLD--PV 362
          :|    ||||   :    ||::||  : |       :| ||    |||: |   :    :
Db    355 FAVDLTMELYQHLESKEWFVQLYYHGKEQ-------VPRGCPDGLCPLDMFLNAMSVYTL 407
```

```
Qy    363 IPQDWATECMAT 374
          |: :   | |
Db    408 SPEKYHALCSQT 419
```

RESULT 13
US-10-121-063-330
; Sequence 330, Application US/10121063
; GENERAL INFORMATION:
; APPLICANT: Baker,Kevin P.
; APPLICANT: Beresini,Maureen
; APPLICANT: DeForge,Laura
; APPLICANT: Desnoyers,Luc
; APPLICANT: Filvaroff,Ellen
; APPLICANT: Gao,Wei-Qiang
; APPLICANT: Gerritsen,Mary E.
; APPLICANT: Goddard,Audrey
; APPLICANT: Godowski,Paul J.
; APPLICANT: Gurney,Austin L.
; APPLICANT: Sherwood,Steven
; APPLICANT: Smith,Victoria
; APPLICANT: Stewart,Timothy A.
; APPLICANT: Tumas,Daniel
; APPLICANT: Watanabe,Colin K
; APPLICANT: Wood,William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3330R1C19
; CURRENT APPLICATION NUMBER: US/10/121,063
; CURRENT FILING DATE: 2002-04-12
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 330
; LENGTH: 428
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-121-063-330

Query Match 15.3%; Score 315.5; DB 6; Length 428;
Best Local Similarity 26.9%; Pred. No. 8.1e-20;
Matches 116; Conservative 62; Mismatches 155; Indels 99; Gaps 20;

```
Qy     16 LGFLLLLSLCLD------------PGQAK------ELKFVTLVFRHGDRGPIETFPTD-- 55
          :| |  |: ||              ||        :|| | :||||| | |::  | :
Db     14 VGVLTSLAYCLHQRRVALAELQEADGQCPVDRSLLKLKMVQVVFRHGARSPLKPLPLEEQ 73

Qy     56 -----------PITESSW--------PQGF-----------------GQLTQWGMEQHY 78
                      | |:  :        |: :                 ||||: ||:| :
Db     74 VEWNPQLLEVPPQTQFDYTVTNLAGGPKPYSPYDSQYHETTLKGGMFAGQLTKVGMQQMF 133

Qy     79 ELGSYIRKRYGR---FLNDTYKHDQIYIRSTDVDRTLMSAMTNLAALF--PPEGISIWNP 133
           ||  :|| |     ||: |:   :::||||:: | | |    || ||    ||
Db    134 ALGERLRKNYVEDIPFLSPTFNPQEVFIRSTNIFRNLESTRCLLAGLFQCQKEG------ 187

Qy    134 RLLWQPIPVHTVSLSEDRLLYLPFRDCPRFEELKSETLESEEFLKRLHPYKSFLDTLSSL 193
               || :||   ::  :|| :: |     |: |    :                | :
Db    188 -----PIIIHT-DEADSEVLYPNYQSC---WSLRQRTRGRRQTASLQPGISEDLKKVKDR 238

Qy    194 SGFDDQDLFGIWSKVYDPLFCESVHNFTLPSWATEDAMIKLKELSELSLLSLYGIHKQ-K 252
           | |  |     :  : | :  |  ||  |||         :: |   :   ||| : |: :
Db    239 MGIDSSDKVDFFI-LLDNVAAEQAHN--LPSCPMLKRFARMIEQRAVD-TSLYILPKEDR 294

Qy    253 EKSRLQGG----VLVNEILKNMKLATQPQKYKKLVMYSAHDTTVSGLQMALDVYNGVLPP 308
          |  ::  |    :| : :|| |  || | | :|| :|:||| |   | | | :::   ||
Db    295 ESLQMAVGPFLHILESNLLKAMDSATAPDKIRKLYLYAAHDVTFIPLLMTLGIFDHKWPP 354

Qy    309 YASCHMMELYH--DKGGHFVEMYYRNETQNEPYPLTLP-GCTHS-CPLEKFAELLD--PV 362
          :|    ||||   :    ||::||  : |       :| ||    |||: |   :    :
Db    355 FAVDLTMELYQHLESKEWFVQLYYHGKEQ-------VPRGCPDGLCPLDMFLNAMSVYTL 407

Qy    363 IPQDWATECMAT 374
```

```
          |: :   |  |
Db    408 SPEKYHALCSQT 419


RESULT  14
US-10-123-108-330
; Sequence 330, Application US/10123108
; GENERAL INFORMATION:
;  APPLICANT: Baker,Kevin P.
;  APPLICANT:  Beresini,Maureen
;  APPLICANT:  DeForge,Laura
;  APPLICANT:  Desnoyers,Luc
;  APPLICANT:  Filvaroff,Ellen
;  APPLICANT:  Gao,Wei-Qiang
;  APPLICANT:  Gerritsen,Mary E.
;  APPLICANT:  Goddard,Audrey
;  APPLICANT:  Godowski,Paul J.
;  APPLICANT:  Gurney,Austin L.
;  APPLICANT:  Sherwood,Steven
;  APPLICANT:  Smith,Victoria
;  APPLICANT:  Stewart,Timothy A.
;  APPLICANT:  Tumas,Daniel
;  APPLICANT:  Watanabe,Colin K
;  APPLICANT:  Wood,William
;  APPLICANT:  Zhang, Zemin
;  TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
;  TITLE OF INVENTION:  ACIDS ENCODING THE SAME
;  FILE REFERENCE: P3330R1C36
;  CURRENT APPLICATION NUMBER: US/10/123,108
;  CURRENT FILING DATE:  2002-04-15
;  PRIOR APPLICATION NUMBER: 60/049911
;  PRIOR FILING DATE: 1997-06-18
;  PRIOR APPLICATION NUMBER: 60/056974
;  PRIOR FILING DATE: 1997-08-26
;  PRIOR APPLICATION NUMBER: 60/059113
;  PRIOR FILING DATE: 1997-09-17
;  PRIOR APPLICATION NUMBER: 60/059115
;  PRIOR FILING DATE: 1997-09-17
;  PRIOR APPLICATION NUMBER: 60/059117
;  PRIOR FILING DATE: 1997-09-17
;  PRIOR APPLICATION NUMBER: 60/059122
;  PRIOR FILING DATE: 1997-09-17
;  PRIOR APPLICATION NUMBER: 60/059184
;  PRIOR FILING DATE: 1997-09-17
;  PRIOR APPLICATION NUMBER: 60/059263
;  PRIOR FILING DATE: 1997-09-18
;  PRIOR APPLICATION NUMBER: 60/059352
;  PRIOR FILING DATE: 1997-09-19
;  PRIOR APPLICATION NUMBER: 60/059588
;  PRIOR FILING DATE: 1997-09-19
;  PRIOR APPLICATION NUMBER: 60/059836
;  PRIOR FILING DATE: 1997-09-24
;  PRIOR APPLICATION NUMBER: 60/062250
;  PRIOR FILING DATE: 1997-10-17
;  PRIOR APPLICATION NUMBER: 60/062285
;  PRIOR FILING DATE: 1997-10-17
;  PRIOR APPLICATION NUMBER: 60/062287
;  PRIOR FILING DATE: 1997-10-17
;  PRIOR APPLICATION NUMBER: 60/062814
;  PRIOR FILING DATE: 1997-10-24
;  PRIOR APPLICATION NUMBER: 60/062816
;  PRIOR FILING DATE: 1997-10-24
;  PRIOR APPLICATION NUMBER: 60/063045
;  PRIOR FILING DATE: 1997-10-24
;  PRIOR APPLICATION NUMBER: 60/063082
;  PRIOR FILING DATE: 1997-10-31
;  PRIOR APPLICATION NUMBER: 60/063127
;  PRIOR FILING DATE: 1997-10-24
;  PRIOR APPLICATION NUMBER: 60/063327
;  PRIOR FILING DATE: 1997-10-27
;  PRIOR APPLICATION NUMBER: 60/063329
;  PRIOR FILING DATE: 1997-10-27
;  PRIOR APPLICATION NUMBER: 60/063550
;  PRIOR FILING DATE: 1997-10-28
;  PRIOR APPLICATION NUMBER: 60/063561
;  PRIOR FILING DATE: 1997-10-28
;  PRIOR APPLICATION NUMBER: 60/063704
;  PRIOR FILING DATE: 1997-10-29
;  PRIOR APPLICATION NUMBER: 60/063733
;  PRIOR FILING DATE: 1997-10-29
;  PRIOR APPLICATION NUMBER: 60/063735
;  PRIOR FILING DATE: 1997-10-29
;  PRIOR APPLICATION NUMBER: 60/063738
;  PRIOR FILING DATE: 1997-10-29
;  PRIOR APPLICATION NUMBER: 60/063755
;  PRIOR FILING DATE: 1997-10-17
;  PRIOR APPLICATION NUMBER: 60/064248
;  PRIOR FILING DATE: 1997-11-03
;  PRIOR APPLICATION NUMBER: 60/064809
;  PRIOR FILING DATE: 1997-11-07
;  PRIOR APPLICATION NUMBER: 60/065186
;  PRIOR FILING DATE: 1997-11-12
;  PRIOR APPLICATION NUMBER: 60/065846
;  PRIOR FILING DATE: 1997-11-17
;  PRIOR APPLICATION NUMBER: 60/066364
;  PRIOR FILING DATE: 1997-11-21
;  PRIOR APPLICATION NUMBER: 60/066453
;  PRIOR FILING DATE: 1997-11-24
;  PRIOR APPLICATION NUMBER: 60/066511
;  PRIOR FILING DATE: 1997-11-24
;  PRIOR APPLICATION NUMBER: 60/066770
;  PRIOR FILING DATE: 1997-11-24
;  PRIOR APPLICATION NUMBER: 60/069212
;  PRIOR FILING DATE: 1997-12-11
;  PRIOR APPLICATION NUMBER: 60/069278
;  PRIOR FILING DATE: 1997-12-11
;  PRIOR APPLICATION NUMBER: 60/069334
;  PRIOR FILING DATE: 1997-12-11
;  PRIOR APPLICATION NUMBER: 60/069694
;  PRIOR FILING DATE: 1997-12-16
;  PRIOR APPLICATION NUMBER: 60/072320
;  PRIOR FILING DATE: 1998-01-23
;  PRIOR APPLICATION NUMBER: 60/073612
;  PRIOR FILING DATE: 1998-02-04
;  PRIOR APPLICATION NUMBER: 60/074086
;  PRIOR FILING DATE: 1998-02-09
;  PRIOR APPLICATION NUMBER: 60/074092
;  PRIOR FILING DATE: 1998-02-09
;  PRIOR APPLICATION NUMBER: 60/077791
;  PRIOR FILING DATE: 1998-03-12
;  PRIOR APPLICATION NUMBER: 60/078910
;  PRIOR FILING DATE: 1998-03-20
;  PRIOR APPLICATION NUMBER: 60/079294
;  PRIOR FILING DATE: 1998-03-25
;  PRIOR APPLICATION NUMBER: 60/079663
;  PRIOR FILING DATE: 1998-02-27
;  PRIOR APPLICATION NUMBER: 60/079728
;  PRIOR FILING DATE: 1998-03-27
;  PRIOR APPLICATION NUMBER: 60/080165
;  PRIOR FILING DATE: 1998-03-31
;  PRIOR APPLICATION NUMBER: 60/081203
;  PRIOR FILING DATE: 1998-04-09
;  PRIOR APPLICATION NUMBER: 60/081229
;  PRIOR FILING DATE: 1998-04-09
;  PRIOR APPLICATION NUMBER: 60/081695
;  PRIOR FILING DATE: 1998-04-14
;  PRIOR APPLICATION NUMBER: 60/081817
;  PRIOR FILING DATE: 1998-04-15
;  PRIOR APPLICATION NUMBER: 60/081818
;  PRIOR FILING DATE: 1998-04-15
;  PRIOR APPLICATION NUMBER: 60/082999
;  PRIOR FILING DATE: 1998-04-24
;  PRIOR APPLICATION NUMBER: 60/083322
;  PRIOR FILING DATE: 1998-04-28
;  PRIOR APPLICATION NUMBER: 60/083545
;  PRIOR FILING DATE: 1998-04-29
```

```
;  PRIOR APPLICATION NUMBER: 60/084600
;  PRIOR FILING DATE: 1998-05-07
;  PRIOR APPLICATION NUMBER: 60/084627
;  PRIOR FILING DATE: 1998-05-07
;  PRIOR APPLICATION NUMBER: 60/084637
;  PRIOR FILING DATE: 1998-05-07
;  PRIOR APPLICATION NUMBER: 60/085149
;  PRIOR FILING DATE: 1998-05-12
;  PRIOR APPLICATION NUMBER: 60/085323
;  PRIOR FILING DATE: 1998-05-13
;  PRIOR APPLICATION NUMBER: 60/085338
;  PRIOR FILING DATE: 1998-05-13
;  PRIOR APPLICATION NUMBER: 60/085339
;  PRIOR FILING DATE: 1998-05-13
;  PRIOR APPLICATION NUMBER: 60/085579
;  PRIOR FILING DATE: 1998-05-15
;  PRIOR APPLICATION NUMBER: 60/085697
;  PRIOR FILING DATE: 1998-05-15
;  PRIOR APPLICATION NUMBER: 60/085704
;  PRIOR FILING DATE: 1998-05-15
;  PRIOR APPLICATION NUMBER: 60/086414
;  PRIOR FILING DATE: 1998-05-22
;  PRIOR APPLICATION NUMBER: 60/086430
;  PRIOR FILING DATE: 1998-05-22
;  PRIOR APPLICATION NUMBER: 60/087106
;  PRIOR FILING DATE: 1998-05-28
;  PRIOR APPLICATION NUMBER: 60/088026
;  PRIOR FILING DATE: 1998-06-04
;  PRIOR APPLICATION NUMBER: 60/088730
;  PRIOR FILING DATE: 1998-06-10
;  PRIOR APPLICATION NUMBER: 60/088741
;  PRIOR FILING DATE: 1998-06-10
;  PRIOR APPLICATION NUMBER: 60/088810
;  PRIOR FILING DATE: 1998-06-10
;  PRIOR APPLICATION NUMBER: 60/088858
;  PRIOR FILING DATE: 19/98-06-11
;  PRIOR APPLICATION NUMBER: 60/089532
;  PRIOR FILING DATE: 1998-06-17
;  PRIOR APPLICATION NUMBER: 60/089599
;  PRIOR FILING DATE: 1998-06-17
;  PRIOR APPLICATION NUMBER: 60/089907
;  PRIOR FILING DATE: 1998-06-18
;  PRIOR APPLICATION NUMBER: 60/089947
;  PRIOR FILING DATE: 1998-06-19
;  PRIOR APPLICATION NUMBER: 60/090349
;  PRIOR FILING DATE: 1998-06-23
;  PRIOR APPLICATION NUMBER: 60/090429
;  PRIOR FILING DATE: 1998-06-24
;  PRIOR APPLICATION NUMBER: 60/090445
;  PRIOR FILING DATE: 1998-06-24
;  PRIOR APPLICATION NUMBER: 60/090538
;  PRIOR FILING DATE: 1998-06-24
;  PRIOR APPLICATION NUMBER: 60/090863
;  PRIOR FILING DATE: 1998-06-26
;  PRIOR APPLICATION NUMBER: 60/091360
;  PRIOR FILING DATE: 1998-07-01
;  PRIOR APPLICATION NUMBER: 60/091519
;  PRIOR FILING DATE: 1998-07-02
;  PRIOR APPLICATION NUMBER: 60/091982
;  PRIOR FILING DATE: 1998-07-07

  Query Match             15.3%;  Score 315.5;  DB 6;  Length 428;
  Best Local Similarity   26.9%;  Pred. No. 8.1e-20;
  Matches  116;  Conservative   62;  Mismatches  155;  Indels   99;  Gaps   20;

Qy      16 LGFLLLLSLCLD------------PGQAK------ELKFVTLVFRHGDRGPIETFPTD-- 55
           :| |  |: ||              ||        :|| | :||||| | |::  | :
Db      14 VGVLTSLAYCLHQRRVALAELQEADGQCPVDRSLLKLKMVQVVFRHGARSPLKPLPLEEQ 73

Qy      56 -----------PITESSW--------PQGF-----------------GQLTQWGMEQHY 78
                      | |:  :        |: :                 ||||: ||:| :
Db      74 VEWNPQLLEVPPQTQFDYTVTNLAGGPKPYSPYDSQYHETTLKGGMFAGQLTKVGMQQMF 133

Qy      79 ELGSYIRKRYGR---FLNDTYKHDQIYIRSTDVDRTLMSAMTNLAALF--PPEGISIWNP 133
           ||  :|| |     ||: |:   :::|||:: | | |    || ||    ||
Db     134 ALGERLRKNYVEDIPFLSPTFNPQEVFIRSTNIFRNLESTRCLLAGLFQCQKEG------ 187

Qy     134 RLLWQPIPVHTVSLSEDRLLYLPFRDCPRFEELKSETLESEEFLKRLHPYKSFLDTLSSL 193
             || :||   ::  :||  :: |     |: |    :           |  :
Db     188 -----PIIIHT-DEADSEVLYPNYQSC---WSLRQRTRGRRQTASLQPGISEDLKKVKDR 238

Qy     194 SGFDDQDLFGIWSKVYDPLFCESVHNFTLPSWATEDAMIKLKELSELSLLSLYGIHKQ-K 252
           | |  |    :  : | :  |  ||  |||        :: |   :   ||| : |: :
Db     239 MGIDSSDKVDFFI-LLDNVAAEQAHN--LPSCPMLKRFARMIEQRAVD-TSLYILPKEDR 294

Qy     253 EKSRLQGG----VLVNEILKNMKLATQPQKYKKLVMYSAHDTTVSGLQMALDVYNGVLPP 308
           |  ::  |    :| : :|| |  || | | :|| :|:||| |   | | | :::   ||
Db     295 ESLQMAVGPFLHILESNLLKAMDSATAPDKIRKLYLYAAHDVTFIPLLMTLGIFDHKWPP 354

Qy     309 YASCHMMELYH--DKGGHFVEMYYRNETQNEPYPLTLP-GCTHS-CPLEKFAELLD--PV 362
           :|    |||  :    ||::||  : |       :| ||    |||: |   :   :
Db     355 FAVDLTMELYQHLESKEWFVQLYYHGKEQ------VPRGCPDGLCPLDMFLNAMSVYTL 407

Qy     363 IPQDWATECMAT 374
           |: :   |  |
Db     408 SPEKYHALCSQT 419


RESULT  15
US-10-123-154-330
; Sequence 330, Application US/10123154
; GENERAL INFORMATION:
;  APPLICANT: Baker,Kevin P.
;  APPLICANT:  Beresini,Maureen
;  APPLICANT:  DeForge,Laura
;  APPLICANT:  Desnoyers,Luc
;  APPLICANT:  Filvaroff,Ellen
;  APPLICANT:  Gao,Wei-Qiang
;  APPLICANT:  Gerritsen,Mary E.
;  APPLICANT:  Goddard,Audrey
;  APPLICANT:  Godowski,Paul J.
;  APPLICANT:  Gurney,Austin L.
;  APPLICANT:  Sherwood,Steven
;  APPLICANT:  Smith,Victoria
;  APPLICANT:  Stewart,Timothy A.
;  APPLICANT:  Tumas,Daniel
;  APPLICANT:  Watanabe,Colin K
;  APPLICANT:  Wood,William
;  APPLICANT:  Zhang, Zemin
;  TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
;  TITLE OF INVENTION:  ACIDS ENCODING THE SAME
;  FILE REFERENCE: P3330R1C39
;  CURRENT APPLICATION NUMBER: US/10/123,154
;  CURRENT FILING DATE:  2002-04-15
; Prior Application removed -  See Palm or File Wrapper
;  NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 330
;   LENGTH: 428
;   TYPE: PRT
;   ORGANISM: Homo Sapien
US-10-123-154-330


  Query Match             15.3%;  Score 315.5;  DB 6;  Length 428;
  Best Local Similarity   26.9%;  Pred. No. 8.1e-20;
  Matches  116;  Conservative   62;  Mismatches  155;  Indels   99;  Gaps   20;

Qy      16 LGFLLLLSLCLD------------PGQAK------ELKFVTLVFRHGDRGPIETFPTD-- 55
           :| |  |: ||              ||        :|| | :||||| | |::  | :
Db      14 VGVLTSLAYCLHQRRVALAELQEADGQCPVDRSLLKLKMVQVVFRHGARSPLKPLPLEEQ 73

Qy      56 -----------PITESSW--------PQGF-----------------GQLTQWGMEQHY 78
                      | |:  :        |: :                 ||||: ||:| :
Db      74 VEWNPQLLEVPPQTQFDYTVTNLAGGPKPYSPYDSQYHETTLKGGMFAGQLTKVGMQQMF 133

Qy      79 ELGSYIRKRYGR---FLNDTYKHDQIYIRSTDVDRTLMSAMTNLAALF--PPEGISIWNP 133
```

```
          ||  :|| |     ||: |:   :::||||:: | | |    || ||    ||
Db    134 ALGERLRKNYVEDIPFLSPTFNPQEVFIRSTNIFRNLESTRCLLAGLFQCQKEG------ 187

Qy    134 RLLWQPIPVHTVSLSEDRLLYLPFRDCPRFEELKSETLESEEFLKRLHPYKSFLDTLSSL 193
            || :||   ::  :||  :: |     |:  |    :              |  :
Db    188 -----PIIIHT-DEADSEVLYPNYQSC---WSLRQRTRGRRQTASLQPGISEDLKKVKDR 238

Qy    194 SGFDDQDLFGIWSKVYDPLFCESVHNFTLPSWATEDAMIKLKELSELSLLSLYGIHKQ-K 252
          | |  |    :  : | :  |  ||  |||        :: |   :   |||  : |: :
Db    239 MGIDSSDKVDFFI-LLDNVAAEQAHN--LPSCPMLKRFARMIEQRAVD-TSLYILPKEDR 294

Qy    253 EKSRLQGG----VLVNEILKNMKLATQPQKYKKLVMYSAHDTTVSGLQMALDVYNGVLPP 308
          |  ::  |    :| : :|| |  || | | :|| :|:||| |   | | | :::   ||
Db    295 ESLQMAVGPFLHILESNLLKAMDSATAPDKIRKLYLYAAHDVTFIPLLMTLGIFDHKWPP 354

Qy    309 YASCHMMELYH--DKGGHFVEMYYRNETQNEPYPLTLP-GCTHS-CPLEKFAELLD--PV 362
          :|    ||||   :    ||::||  : |       :| ||    |||: |   :    :
Db    355 FAVDLTMELYQHLESKEWFVQLYYHGKEQ-------VPRGCPDGLCPLDMFLNAMSVYTL 407

Qy    363 IPQDWATECMAT 374
          |: :   |  |
Db    408 SPEKYHALCSQT 419


Search completed: June  4, 2002, 10:31:38
Job time: 109 sec
```

GenCore version 4.5

OM protein - protein search, using sw model

Run on: June 4, 2002, 10:29:49 ; Search time 31.95 Seconds
(without alignments)
1338.448 Million cell updates/sec

Title: US-09-402-845-2
Perfect score: 2060
Sequence: 1 MGAVPLPLSPTASLSLGFLL.....DWATECMATSSHQGTVGALG 385

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_032802:*
1: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1980.DAT:*
2: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1981.DAT:*
3: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1982.DAT:*
4: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1983.DAT:*
5: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1984.DAT:*
6: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1985.DAT:*
7: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1986.DAT:*
8: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1987.DAT:*
9: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1988.DAT:*
10: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1989.DAT:*
11: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1990.DAT:*
12: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1991.DAT:*
13: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1992.DAT:*
14: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1993.DAT:*
15: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1994.DAT:*
16: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1995.DAT:*
17: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1996.DAT:*
18: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1997.DAT:*
19: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1998.DAT:*
20: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1999.DAT:*
21: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA2000.DAT:*
22: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	2060	100.0	385	20	AAW30574	Mouse prostatic ac
2	1661.5	80.7	386	19	AAW57418	Protein encoded by
3	1661.5	80.7	386	20	AAW95081	Protein sequence S
4	1661.5	80.7	386	21	AAY59293	Prostatic acid pho
5	1661.5	80.7	386	22	AAU06277	Prostatic Acid Pho
6	1661.5	80.7	386	22	AAU02172	Biomarker UC band
7	1661.5	80.7	386	22	AAG62145	Human prostatic ac
8	1661.5	80.7	386	22	AAB74820	Prostate tumour an
9	1661.5	80.7	515	18	AAW19762	PAP-GM-CSF immunos
10	939	45.6	216	21	AAB56451	Human prostate can
11	880.5	42.7	210	21	AAG00216	Human secreted pro
12	745.5	36.2	426	22	ABG23348	Novel human diagno
13	689.5	33.5	178	21	AAG00217	Human secreted pro
14	669.5	32.5	171	21	AAB56447	Human prostate can
15	616	29.9	438	22	ABB58803	Drosophila melanog
16	567	27.5	392	22	ABB71760	Drosophila melanog
17	487	23.6	410	22	ABB71761	Drosophila melanog
18	476	23.1	412	22	ABB71762	Drosophila melanog
19	388.5	18.9	223	22	ABG04588	Novel human diagno
20	384.5	18.7	395	22	ABB62558	Drosophila melanog
21	315.5	15.3	421	21	AAB09969	Human lysophosphat
22	315.5	15.3	428	20	AAY13368	Amino acid sequenc
23	315.5	15.3	428	21	AAB24399	Human PRO231 prote
24	315.5	15.3	428	22	AAM38841	Human polypeptide
25	315.5	15.3	428	22	AAU12336	Human PRO231 polyp
26	315.5	15.3	428	22	AAE05764	Human secreted pro
27	315.5	15.3	470	22	AAM40627	Human polypeptide
28	313.5	15.2	428	22	AAB80236	Human PRO231 prote
29	260.5	12.6	106	21	AAG03770	Human secreted pro
30	208.5	10.1	480	20	AAY41742	Human PRO706 prote
31	208.5	10.1	480	21	AAB44298	Human PRO706 (UNQ3
32	208.5	10.1	480	22	AAU29064	Human PRO polypept
33	208.5	10.1	480	22	AAB87346	Human gene 5 encod
34	183	8.9	402	22	ABG11885	Novel human diagno
35	175	8.5	519	22	AAB31601	Amino acid sequenc
36	169	8.2	468	21	AAG30243	Arabidopsis thalia
37	164.5	8.0	398	17	AAR94685	Flea saliva protei
38	157	7.6	213	22	AAM91922	Human digestive sy
39	154	7.5	31	20	AAW30575	Mouse prostatic ac
40	153.5	7.5	141	22	AAB20323	Human protein phos
41	151.5	7.4	353	18	AAW30484	Flea saliva protei
42	150.5	7.3	353	19	AAW82374	Flea saliva protei
43	147.5	7.2	524	22	AAB31600	Amino acid sequenc
44	144.5	7.0	304	22	ABG09388	Novel human diagno
45	144.5	7.0	375	19	AAW82382	Flea saliva protei

ALIGNMENTS

RESULT 1
AAW30574
ID AAW30574 standard; Protein; 385 AA.
XX
AC AAW30574;
XX
DT 01-MAR-1999 (first entry)
XX
DE Mouse prostatic acid phosphatase.
XX
KW Prostatic acid phosphatase; PAP; mouse; tumour related antigen;
KW diagnosis; vaccine.
XX
OS Mus sp.
XX
FH Key Location/Qualifiers
FT Peptide 1..31
FT /label= Sig_peptide
FT Misc-difference 364
FT /note= "encoded by CCB"
XX
PN WO9846769-A1.
XX
PD 22-OCT-1998.
XX
PF 10-APR-1998; 98WO-US07232.
XX
PR 11-APR-1997; 97US-0043301.
XX
PA (DEND-) DENDREON CORP.
XX
PI Laus R, Ruegg CL, Shapero MH, Yang D;
XX
DR WPI; 1999-009335/01.

```
DR   N-PSDB; AAV45592.
XX
PT   New mouse prostatic acid phosphatase - used to induce an immune
PT   response against tumour related antigens
XX
PS   Claim 1; Page 20-21; 30pp; English.
XX
CC   This is the amino acid sequence of mouse prostatic acid phosphatase
CC   (PAP), a novel tumour associated antigen.  PAP cDNA (see AAV45592)
CC   was cloned from mouse prostate organ cDNA, and can be used in the
CC   recombinant production of mouse PAP.  A method for producing an
CC   immune response against an autologous polypeptide tumour antigen
CC   (e.g. human PAP) involves immunising a subject with a xenogeneic
CC   antigen (e.g. mouse PAP), either alone, as part of a viral antigen
CC   construct, or as part of a pulsed dendritic cell preparation.
CC   Recombinant viruses expressing PAP are used in compositions to
CC   elicit immune responses against a tumour related antigen.  The
CC   compositions are also useful for reducing tumour cell load.
XX
SQ   Sequence   385 AA;


  Query Match           100.0%;  Score 2060;  DB 20;  Length 385;
  Best Local Similarity  100.0%;  Pred. No. 7.2e-196;
  Matches  385;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 MGAVPLPLSPTASLSLGFLLLLSLCLDPGQAKELKFVTLVFRHGDRGPIETFPTDPITES 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 mgavplplsptaslslgfllllslcldpgqakelkfvtlvfrhgdrgpietfptdpites 60

Qy      61 SWPQGFGQLTQWGMEQHYELGSYIRKRYGRFLNDTYKHDQIYIRSTDVDRTLMSAMTNLA 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 swpqgfgqltqwgmeqhyelgsyirkrygrflndtykhdqiyirstdvdrtlmsamtnla 120

Qy     121 ALFPPEGISIWNPRLLWQPIPVHTVSLSEDRLLYLPFRDCPRFEELKSETLESEEFLKRL 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 alfppegisiwnprllwqpipvhtvslsedrllylpfrdcprfeelksetleseeflkrl 180

Qy     181 HPYKSFLDTLSSLSGFDDQDLFGIWSKVYDPLFCESVHNFTLPSWATEDAMIKLKELSEL 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 hpyksfldtlsslsgfddqdlfgiwskvydplfcesvhnftlpswatedamiklkelsel 240

Qy     241 SLLSLYGIHKQKEKSRLQGGVLVNEILKNMKLATQPQKYKKLVMYSAHDTTVSGLQMALD 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 sllslygihkqkeksrlqggvlvneilknmklatqpqkykklvmysahdttvsglqmald 300

Qy     301 VYNGVLPPYASCHMMELYHDKGGHFVEMYYRNETQNEPYPLTLPGCTHSCPLEKFAELLD 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 vyngvlppyaschmmelyhdkgghfvemyyrnetqnepypltlpgcthscplekfaelld 360

Qy     361 PVIPQDWATECMATSSHQGTVGALG 385
           |||||||||||||||||||||||||
Db     361 pvipqdwatecmatsshqgtvgalg 385


RESULT   2
AAW57418
ID   AAW57418 standard; Protein; 386 AA.
XX
AC   AAW57418;
XX
DT   07-AUG-1998  (first entry)
XX
DE   Protein encoded by a human prostate cancer marker.
XX
KW   Prostate cancer; human; marker; diagnosis; treatment; probe.
XX
OS   Homo sapiens.
XX
PN   WO9804689-A1.
XX
PD   05-FEB-1998.
XX
PF   31-JUL-1996;   96WO-US12516.
XX
PR   31-JUL-1996;   96WO-US12516.
XX
PA   (UROC-) UROCOR INC.
XX
PI   An G,  O'hara SM,  Ralph D,  Veltri R;
XX
DR   WPI; 1998-130681/12.
DR   N-PSDB; AAV29653.
XX
PT   Human prostate cancer marker - useful for detection and treatment of
PT   human prostate cancer
XX
PS   Disclosure; Pages 167-174; 229pp; English.
XX
CC   This protein is encoded by a marker sequence for human prostate cancer.
CC   Isolated nucleic acid segments shown in AAV16881 to AAV16885, AAV16890
CC   to AAV16903, AAV26351 and AAV26352 which can act as human prostate cancer
CC   markers are provided in the specification. It also provides methods for
CC   identifying markers for human prostate cancer and for detection of
CC   prostate cancer cells. The markers can be identified by amplifying human
CC   prostate RNA to provide nucleic acid amplification products, separating
CC   the products and identifying those RNA that are differentially expressed
CC   between human prostate cancers versus normal or benign human prostate.
CC   Prostate cancer cells in a sample can be detected by detecting a nucleic
CC   acid in a sample, the nucleic acid being a prostate cancer marker.
CC   Probes and primers derived from this marker can be used for the detection
CC   of prostate cancer cells in a sample. Antibodies against the protein
CC   encoded by the marker nucleic acid fragments, inhibitors of the protein
CC   and oligonucleotides antisense to the markers can be used in the
CC   treatment of prostate cancer. The antibodies can also be used for the
CC   diagnosis of human prostate cancer.
XX
SQ   Sequence   386 AA;


  Query Match           80.7%;  Score 1661.5;  DB 19;  Length 386;
  Best Local Similarity  81.6%;  Pred. No. 2.9e-156;
  Matches  311;  Conservative   28;  Mismatches   41;  Indels    1;  Gaps    1;

Qy       1 MGAVPLPLSPTASLSLGFLLLLSLCLDPG-QAKELKFVTLVFRHGDRGPIETFPTDPITE 59
           | | || |:  |||||||| ||   ||    ||||||||||||||| ||:||||||| |
Db       1 mraaplllaraaslslgflfllffwldrsvlakelkfvtlvfrhgdrspidtfptdpike 60

Qy      60 SSWPQGFGQLTQWGMEQHYELGSYIRKRYGRFLNDTYKHDQIYIRSTDVDRTLMSAMTNL 119
           |||||||||||| ||||||||| |||||| :|||::|||:|:|||||||||||||||||
Db      61 sswpqgfgqltqlgmeqhyelgeyirkryrkflnesykheqvyirstdvdrtlmsamtnl 120

Qy     120 AALFPPEGISIWNPRLLWQPIPVHTVSLSEDRLLYLPFRDCPRFEELKSETLESEEFLKR 179
           ||||||||:||||| |||||||||| ||||:|||||||:||||:||:||||:|||| ||
Db     121 aalfppegvsiwnpillwqpipvhtvplsedqllylpfrncprfqelesetlkseefqkr 180

Qy     180 LHPYKSFLDTLSSLSGFDDQDLFGIWSKVYDPLFCESVHNFTLPSWATEDAMIKLKELSE 239
           ||||| |: ||  |||   |||||||||||||:|||||||||||||| | ||:||||
Db     181 lhpykdfiatlgklsglhgqdlfgiwskvydplycesvhnftlpswatedtmtklrelse 240

Qy     240 LSLLSLYGIHKQKEKSRLQGGVLVNEILKNMKLATQPQKYKKLVMYSAHDTTVSGLQMAL 299
           |||||||||||||||||||||||||| :|| |||   ||||:|||||||||||||||
Db     241 lsllslygihkqkeksrlqggvlvneilnhmkratqipsykklimysahdttvsglqmal 300

Qy     300 DVYNGVLPPYASCHMMELYHDKGGHFVEMYYRNETQNEPYPLTLPGCTHSCPLEKFAELL 359
           |||||:|||||||: ||| :|| :||||||||||:|||||  |||| |||||:||||:
Db     301 dvyngllppyaschltelyfekgeyfvemyyrnetqhepyplmlpgcspscplerfaelv 360

Qy     360 DPVIPQDWATECMATSSHQGT 380
           |||||||:|||| |:|||||
Db     361 gpvipqdwstecmttnshqgt 381


RESULT   3
AAW95081
ID   AAW95081 standard; Protein; 386 AA.
```

```
XX
AC   AAW95081;
XX
DT   20-MAY-1999  (first entry)
XX
DE   Protein sequence Seq ID No: 47 from US 5882864.
XX
KW   Prostate cancer; benign prostatic hyperplasia; marker gene; tumour;
KW   differentiation; Reverse Transcription Polymerase Chain Reaction;
KW   diagnostic; progression; cancer; metastasis; RT-PCR.
XX
OS   Homo sapiens.
XX
PN   US5882864-A.
XX
PD   16-MAR-1999.
XX
PF   31-JUL-1996;   96US-0692787.
XX
PR   31-JUL-1995;   95US-0001655.
PR   31-JUL-1996;   96US-0692787.
XX
PA   (UROC-) UROCOR INC.
XX
PI   An G,  O'Hara SM,  Ralph D,  Veltri R;
XX
DR   WPI; 1999-214055/18.
DR   N-PSDB; AAX26062.
XX
PT   Diagnosing prostate cancer and benign prostatic hyperplasia cells -
PT   using oligonucleotide probes specific for marker genes associated
PT   with tumor differentiation and progression in Reverse Transcription
PT   Polymerase Chain Reaction analysis
XX
PS   Disclosure; Column 93-98; 74pp; English.
XX
CC   The invention relates to methods for diagnosing prostate cancer or benign
CC   prostatic hyperplasia cells in a biological sample. The method uses
CC   oligonucleotides specific for marker genes associated with tumour
CC   differentiation and progression in Reverse Transcription Polymerase Chain
CC   Reaction (RT-PCR) analysis. The methods are diagnostic techniques useful
CC   for detecting and monitoring the progression of benign prostatic
CC   hyperplasia and human prostate cancer (the most prevalent form of cancer
CC   and a major cause of death in males) prior to the tumor undergoing
CC   metastasis, therefore allowing the optimal method of treatment to be
CC   determined before the condition becomes life threatening.
XX
SQ   Sequence   386 AA;

  Query Match             80.7%;  Score 1661.5;  DB 20;  Length 386;
  Best Local Similarity   81.6%;  Pred. No. 2.9e-156;
  Matches  311;  Conservative   28;  Mismatches   41;  Indels    1;  Gaps    1;

Qy       1 MGAVPLPLSPTASLSLGFLLLLSLCLDPG-QAKELKFVTLVFRHGDRGPIETFPTDPITE 59
           | | || |:  |||||||| ||   ||    ||||||||||||||| ||:||||||| |
Db       1 mraaplllaraaslslgflfllffwldrsvlakelkfvtlvfrhgdrspidtfptdpike 60

Qy      60 SSWPQGFGQLTQWGMEQHYELGSYIRKRYGRFLNDTYKHDQIYIRSTDVDRTLMSAMTNL 119
           |||||||||||| |||||||| |||||| :|||::|||:|:|||||||||||||||||||
Db      61 sswpqgfgqltqlgmeqhyelgeyirkryrkflnesykheqvyirstdvdrtlmsamtnl 120

Qy     120 AALFPPEGISIWNPRLLWQPIPVHTVSLSEDRLLYLPFRDCPRFEELKSETLESEEFLKR 179
           ||||||||:||||| |||||||||||| ||||:|||||||:||||:||:||||:|||| ||
Db     121 aalfppegvsiwnpillwqpipvhtvplsedqllylpfrncprfqelesetlkseefqkr 180

Qy     180 LHPYKSFLDTLSSLSGFDDQDLFGIWSKVYDPLFCESVHNFTLPSWATEDAMIKLKELSE 239
           ||||| |: ||  |||   |||||||||||||:|||||||||||||||| | ||:||||
Db     181 lhpykdfiatlgklsglhgqdlfgiwskvydplycesvhnftlpswatedtmtklrelse 240

Qy     240 LSLLSLYGIHKQKEKSRLQGGVLVNEILKNMKLATQPQKYKKLVMYSAHDTTVSGLQMAL 299
           ||||||||||||||||||||||||||| :|| |||   ||||:|||||||||||||||||
Db     241 lsllslygihkqkeksrlqggvlvneilnhmkratqipsykklimysahdttvsglqmal 300

Qy     300 DVYNGVLPPYASCHMMELYHDKGGHFVEMYYRNETQNEPYPLTLPGCTHSCPLEKFAELL 359
           |||||:|||||||: ||| :|| :||||||||||:||||| |||| |||||:||||:
Db     301 dvyngllppyaschltelyfekgeyfvemyyrnetqhepyplmlpgcspscplerfaelv 360

Qy     360 DPVIPQDWATECMATSSHQGT 380
           |||||||:|||| |:|||||
Db     361 gpvipqdwstecmttnshqgt 381


RESULT   4
AAY59293
ID   AAY59293 standard; peptide; 386 AA.
XX
AC   AAY59293;
XX
DT   19-APR-2000  (first entry)
XX
DE   Prostatic acid phosphatase marker UC Band #47 amino acid sequence.
XX
KW   Nucleic acid marker; biomarker; tumour; prostate cancer; bladder cancer;
KW   benign prostatic hyperplasia; BPH; breast cancer; human; immunodetection.
XX
OS   Homo sapiens.
XX
PN   WO9964631-A1.
XX
PD   16-DEC-1999.
XX
PF   11-JUN-1999;   99WO-US13151.
XX
PR   12-JUN-1998;   98US-0097199.
XX
PA   (UROC-) UROCOR INC.
XX
PI   An G,  O'Hara SM,  Ralph D,  Veltri RW;
XX
DR   WPI; 2000-116557/10.
DR   N-PSDB; AAZ87547.
XX
PT   Novel RNA biomarkers for diagnosis, prognosis and management of
PT   prostate, breast and bladder cancer -
XX
PS   Example 2; Page 173-176; 191pp; English.
XX
CC   The invention provides nucleic acid markers of prostate, breast and
CC   bladder cancer. The markers are indicators of malignant transformation of
CC   prostate, breast and bladder tissues and are diagnostic of the potential
CC   for metastatic spread of malignant prostate tumours. The nucleic acid can
CC   also be used as targets for therapeutic intervention in prostate cancer,
CC   benign prostatic hyperplasia (BPH), bladder cancer or breast cancer. The
CC   markers may be used to design specific probes and primers, for the rapid
CC   analysis of prostate, bladder or breast biopsy samples. The probes and
CC   primers may also be used for in situ hybridization or in situ PCR
CC   detection and diagnosis. They may also be used to identify and isolate
CC   full length gene sequences form various DNA libraries. Antibodies
CC   against the polypeptide products of the markers can be used to treat
CC   prostate cancer, bladder cancer or breast cancer. The encoded proteins
CC   may be used to detect antibodies. The proteins and antibodies can be
CC   used in immunodetection methods for detecting or quantifying the cancers,
CC   and for clinical diagnosis of these cancers. The antibodies may also be
CC   used for radioimaging to quantify and localize the encoded proteins..
XX
SQ   Sequence   386 AA;

  Query Match             80.7%;  Score 1661.5;  DB 21;  Length 386;
  Best Local Similarity   81.6%;  Pred. No. 2.9e-156;
  Matches  311;  Conservative   28;  Mismatches   41;  Indels    1;  Gaps    1;

Qy       1 MGAVPLPLSPTASLSLGFLLLLSLCLDPG-QAKELKFVTLVFRHGDRGPIETFPTDPITE 59
           | | || |:  |||||||| ||   ||    ||||||||||||||| ||:||||||| |
Db       1 mraaplllaraaslslgflfllffwldrsvlakelkfvtlvfrhgdrspidtfptdpike 60
```

```
Qy      60 SSWPQGFGQLTQWGMEQHYELGSYIRKRYGRFLNDTYKHDQIYIRSTDVDRTLMSAMTNL 119
           ||||||||||||| |||||||||| |||||| :|||::|||:|:|||||||||||||||||
Db      61 sswpqgfgqltqlgmeqhyelgeyirkryrkflnesykheqvyirstdvdrtlmsamtnl 120

Qy     120 AALFPPEGISIWNPRLLWQPIPVHTVSLSEDRLLYLPFRDCPRFEELKSETLESEEFLKR 179
           |||||||:||||| |||||||||| |||||||||:|||||:||:||||:|||| ||
Db     121 aalfppegvsiwnpillwqpipvhtvplsedqllylpfrncprfqelesetlkseefqkr 180

Qy     180 LHPYKSFLDTLSSLSGFDDQDLFGIWSKVYDPLFCESVHNFTLPSWATEDAMIKLKELSE 239
           ||||| |: ||  |||   |||||||||||||:||||||||||||||| | ||:||||
Db     181 lhpykdfiatlgklsglhgqdlfgiwskvydplycesvhnftlpswatedtmtklrelse 240

Qy     240 LSLLSLYGIHKQKEKSRLQGGVLVNEILKNMKLATQPQKYKKLVMYSAHDTTVSGLQMAL 299
           ||||||||||||||||||||||||||||| :|| |||   ||||:|||||||||||||||
Db     241 lsllslygihkqkeksrlqggvlvneilnhmkratqipsykklimysahdttvsglqmal 300

Qy     300 DVYNGVLPPYASCHMMELYHDKGGHFVEMYYRNETQNEPYPLTLPGCTHSCPLEKFAELL 359
           |||||:|||||||: ||| :|| :|||||||||||:||||| |||||: |||||:||||:
Db     301 dvyngllppyaschltelyfekgeyfvemyyrnetqhepyplmlpgcspscplerfaelv 360

Qy     360 DPVIPQDWATECMATSSHQGT 380
           |||||||:|||| |:|||||
Db     361 gpvipqdwstecmttnshqgt 381


RESULT   5
AAU06277
ID   AAU06277 standard; Protein; 386 AA.
XX
AC   AAU06277;
XX
DT   24-OCT-2001  (first entry)
XX
DE   Prostatic Acid Phosphatase (PAP) polypeptide.
XX
KW   Prostate cancer-associated antigen; supermotif; human leukocyte antigen;
KW   HLA; cytotoxic T lymphocyte; helper T lymphocyte; HLA superfamily; CTL;
KW   immunogenicity; immunosuppression; HTL.
XX
OS   Homo sapiens.
XX
PN   WO200145728-A2.
XX
PD   28-JUN-2001.
XX
PF   20-DEC-2000; 2000WO-US35516.
XX
PR   21-DEC-1999;   99US-0171312.
PR   07-AUG-2000; 2000US-0633364.
XX
PA   (EPIM-) EPIMMUNE INC.
XX
PI   Fikes J,  Sette A,  Sidney J,  Southwood S,  Chesnut R,  Celis E;
PI   Keogh E;
XX
DR   WPI; 2001-398311/42.
XX
PT   Tumour antigen-associated group-based vaccines useful for vaccinating
PT   against prostate cancer -
XX
PS   Example 2; Page 15; 252pp; English.
XX
CC   The sequences represent prostate cancer-associated antigens and derived
CC   motif or supermotif epitopes. The peptide epitopes are included in
CC   prostate cancer vaccine compositions due to their ability to bind to
CC   human leukocyte antigen (HLA) molecules, which recognise the motifs.
CC   Peptides with a high binding affinity are further tested for their
CC   ability to induce a cytotoxic T lymphocyte (CTL) or a helper T lymphocyte
CC   (HTL) response. Supermotif-bearing peptides may also be tested for their
CC   binding affinity to multiple alleles within the HLA superfamily. The
CC   vaccine compositions can be modified, for example, to enhance
CC   immunogenicity, to avoid the inclusion of immunosuppressive groups, or to
CC   alter the immune response to suit the target disease. These group-based
CC   vaccines allow the focus of an immune response to multiple selected
CC   antigens from the same pathogen. Variability among the immune responses
CC   of patients can therefore be alleviated by the inclusion of groups from
CC   multiple antigens in a vaccine.
XX
SQ   Sequence   386 AA;


  Query Match             80.7%;  Score 1661.5;  DB 22;  Length 386;
  Best Local Similarity   81.6%;  Pred. No. 2.9e-156;
  Matches  311;  Conservative   28;  Mismatches   41;  Indels    1;  Gaps    1;

Qy       1 MGAVPLPLSPTASLSLGFLLLLSLCLDPG-QAKELKFVTLVFRHGDRGPIETFPTDPITE 59
           | | || |:  |||||||| ||   ||    |||||||||||||||| ||:||||||| |
Db       1 mraaplllaraaslslgflflllffwldrsvlakelkfvtlvfrhgdrspidtfptdpike 60

Qy      60 SSWPQGFGQLTQWGMEQHYELGSYIRKRYGRFLNDTYKHDQIYIRSTDVDRTLMSAMTNL 119
           |||||||||||| |||||||| |||||| :|||::|||:|:|||||||||||||||||||
Db      61 sswpqgfgqltqlgmeqhyelgeyirkryrkflnesykheqvyirstdvdrtlmsamtnl 120

Qy     120 AALFPPEGISIWNPRLLWQPIPVHTVSLSEDRLLYLPFRDCPRFEELKSETLESEEFLKR 179
           ||||||||:||||| |||||||||| |||||||||:|||||:||:||||:|||| ||
Db     121 aalfppegvsiwnpillwqpipvhtvplsedqllylpfrncprfqelesetlkseefqkr 180

Qy     180 LHPYKSFLDTLSSLSGFDDQDLFGIWSKVYDPLFCESVHNFTLPSWATEDAMIKLKELSE 239
           ||||| |: ||  |||   |||||||||||||:||||||||||||||| | ||:||||
Db     181 lhpykdfiatlgklsglhgqdlfgiwskvydplycesvhnftlpswatedtmtklrelse 240

Qy     240 LSLLSLYGIHKQKEKSRLQGGVLVNEILKNMKLATQPQKYKKLVMYSAHDTTVSGLQMAL 299
           ||||||||||||||||||||||||||||| :|| |||   ||||:|||||||||||||||
Db     241 lsllslygihkqkeksrlqggvlvneilnhmkratqipsykklimysahdttvsglqmal 300

Qy     300 DVYNGVLPPYASCHMMELYHDKGGHFVEMYYRNETQNEPYPLTLPGCTHSCPLEKFAELL 359
           |||||:|||||||: ||| :|| :|||||||||||:||||| |||||: |||||:||||:
Db     301 dvyngllppyaschltelyfekgeyfvemyyrnetqhepyplmlpgcspscplerfaelv 360

Qy     360 DPVIPQDWATECMATSSHQGT 380
           |||||||:|||| |:|||||
Db     361 gpvipqdwstecmttnshqgt 381


RESULT   6
AAU02172
ID   AAU02172 standard; Protein; 386 AA.
XX
AC   AAU02172;
XX
DT   29-AUG-2001  (first entry)
XX
DE   Biomarker UC band 47 (PAP), used in diagnosis and prognosis of cancer.
XX
KW   Prostate; breast; bladder; cancer; biomarker; probe; diagnostic;
KW   benign prostatic hyperplasia; BPH; therapeutic; human;
KW   prostatic acid phosphatase; PAP.
XX
OS   Homo sapiens.
XX
PN   US6218529-B1.
XX
PD   17-APR-2001.
XX
PF   12-JUN-1998;   98US-0097199.
XX
PR   31-JUL-1995;   95US-0001655.
PR   11-JAN-1996;   96US-0013611.
PR   31-JUL-1996;   96US-0692787.
XX
PA   (UROC-) UROCOR INC.
XX
PI   An G,  O'Hara SM,  Ralph D,  Veltri R;
XX
DR   WPI; 2001-289849/30.
```

```
DR   N-PSDB; AAS03766.
XX
PT   New nucleic acids as biomarkers and targets useful for detecting,
PT   diagnosing, prognosing, and in developing treatments for prostate,
PT   breast and bladder cancer -
XX
PS   Disclosure; Column 105; 78pp; English.
XX
CC   The sequence represents the amino acid sequence of biomarker UC band 47,
CC   prostatic acid phosphatase (PAP), used in detection of prostate, breast
CC   and bladder cancer. Biomarker nucleic acid sequences can be used as
CC   hybridisation probes and primers that specifically hybridise to prostate
CC   cancer, benign prostatic hyperplasia (BPH), bladder cancer or breast
CC   cancer markers. Proteins encoded by the nucleic acid markers can be used
CC   to produce antibodies for the detection of prostate, breast or bladder
CC   cancer. The nucleic acids can be used as targets for therapeutic
CC   intervention in these diseases, in the identification and isolation of
CC   full-length gene sequences, including regulatory elements for gene
CC   expression, from genomic human DNA libraries, as hybridisation probes for
CC   screening genomic human DNA libraries. The kits comprising the nucleic
CC   acid sequences are useful for detecting bladder, breast or prostate
CC   cancer cells in a biological sample.
XX
SQ   Sequence   386 AA;


  Query Match             80.7%;  Score 1661.5;  DB 22;  Length 386;
  Best Local Similarity   81.6%;  Pred. No. 2.9e-156;
  Matches  311;  Conservative   28;  Mismatches   41;  Indels    1;  Gaps    1;

Qy       1 MGAVPLPLSPTASLSLGFLLLLSLCLDPG-QAKELKFVTLVFRHGDRGPIETFPTDPITE 59
           | | || |:  |||||||| ||   ||    ||||||||||||||| ||:||||||| |
Db       1 mraaplllaraaslslgflfllffwldrsvlakelkfvtlvfrhgdrspidtfptdpike 60

Qy      60 SSWPQGFGQLTQWGMEQHYELGSYIRKRYGRFLNDTYKHDQIYIRSTDVDRTLMSAMTNL 119
           ||||||||||| |||||||| |||||| :|||::|||:|:|||||||||||||||||||
Db      61 sswpqgfgqltqlgmeqhyelgeyirkryrkflnesykheqvyirstdvdrtlmsamtnl 120

Qy     120 AALFPPEGISIWNPRLLWQPIPVHTVSLSEDRLLYLPFRDCPRFEELKSETLESEEFLKR 179
           ||||||||:||||| |||||||||| |||||:|||||||:||||:||:||||:|||| ||
Db     121 aalfppegvsiwnpillwqpipvhtvplsedqllylpfrncprfqelesetlkseefqkr 180

Qy     180 LHPYKSFLDTLSSLSGFDDQDLFGIWSKVYDPLFCESVHNFTLPSWATEDAMIKLKELSE 239
           ||||| |: ||  |||   |||||||||||||||:||||||||||||||| | ||:||||
Db     181 lhpykdfiatlgklsglhgqdlfgiwskvydplycesvhnftlpswatedtmtklrelse 240

Qy     240 LSLLSLYGIHKQKEKSRLQGGVLVNEILKNMKLATQPQKYKKLVMYSAHDTTVSGLQMAL 299
           |||||||||||||||||||||||||||| :|| |||   ||||:|||||||||||||||
Db     241 lsllslygihkqkeksrlqggvlvneilnhmkratqipsykklimysahdttvsglqmal 300

Qy     300 DVYNGVLPPYASCHMMELYHDKGGHFVEMYYRNETQNEPYPLTLPGCTHSCPLEKFAELL 359
           |||||:||||||||: ||| :|| :|||||||||||:||||| |||||: |||||:||||:
Db     301 dvyngllppyaschltelyfekgeyfvemyyrnetqhepyplmlpgcspscplerfaelv 360

Qy     360 DPVIPQDWATECMATSSHQGT 380
            |||||||:|||| |:|||||
Db     361 gpvipqdwstecmttnshqgt 381


RESULT   7
AAG62145
ID   AAG62145 standard; Protein; 386 AA.
XX
AC   AAG62145;
XX
DT   06-JUL-2001  (first entry)
XX
DE   Human prostatic acid phosphatase SEQ ID NO: 328.
XX
KW   Human; mouse; immunotherapy; cancer; leukaemia; WT1; Wilm's tumour gene;
KW   chromosome 11p13; zinc finger transcription factor.
XX
OS   Homo sapiens.
XX
PN   WO200125273-A2.
XX
PD   12-APR-2001.
XX
PF   04-OCT-2000; 2000WO-US27465.
XX
PR   04-OCT-1999;   99US-0157459.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Skeiky YAW,  Xu J,  Cheever MA,  Reed SG;
XX
DR   WPI; 2001-328324/34.
XX
PT   Polypeptide comprising part of the Wilm's Tumour gene product sequence is
PT   used in the diagnosis and treatment of malignant diseases e.g. leukemia
PT   and cancer associated with WTI -
XX
PS   Disclosure; Page 209-210; 228pp; English.
XX
CC   The present invention describes compositions comprising peptides derived
CC   from the Wilm's tumour protein WT1 and methods for their use in treating
CC   malignant diseases. Peptides derived from both the murine and human WT1
CC   proteins are provided. The human WT1 gene is found on chromosome 11p13,
CC   and the protein was shown to be a zinc finger transcription factor. The
CC   immunogenic peptides of the invention are particularly useful in the
CC   diagnosis and treatment of cancer and leukaemia. The present sequence is
CC   a polypeptide described in the exemplification of the invention.
XX
SQ   Sequence   386 AA;


  Query Match             80.7%;  Score 1661.5;  DB 22;  Length 386;
  Best Local Similarity   81.6%;  Pred. No. 2.9e-156;
  Matches  311;  Conservative   28;  Mismatches   41;  Indels    1;  Gaps    1;

Qy       1 MGAVPLPLSPTASLSLGFLLLLSLCLDPG-QAKELKFVTLVFRHGDRGPIETFPTDPITE 59
           | | || |:  |||||||| ||   ||    ||||||||||||||| ||:||||||| |
Db       1 mraaplllaraaslslgflfllffwldrsvlakelkfvtlvfrhgdrspidtfptdpike 60

Qy      60 SSWPQGFGQLTQWGMEQHYELGSYIRKRYGRFLNDTYKHDQIYIRSTDVDRTLMSAMTNL 119
           ||||||||||| |||||||| |||||| :|||::|||:|:|||||||||||||||||||
Db      61 sswpqgfgqltqlgmeqhyelgeyirkryrkflnesykheqvyirstdvdrtlmsamtnl 120

Qy     120 AALFPPEGISIWNPRLLWQPIPVHTVSLSEDRLLYLPFRDCPRFEELKSETLESEEFLKR 179
           ||||||||:||||| |||||||||| |||||:|||||||:||||:||:||||:|||| ||
Db     121 aalfppegvsiwnpillwqpipvhtvplsedqllylpfrncprfqelesetlkseefqkr 180

Qy     180 LHPYKSFLDTLSSLSGFDDQDLFGIWSKVYDPLFCESVHNFTLPSWATEDAMIKLKELSE 239
           ||||| |: ||  |||   |||||||||||||||:||||||||||||||| | ||:||||
Db     181 lhpykdfiatlgklsglhgqdlfgiwskvydplycesvhnftlpswatedtmtklrelse 240

Qy     240 LSLLSLYGIHKQKEKSRLQGGVLVNEILKNMKLATQPQKYKKLVMYSAHDTTVSGLQMAL 299
           |||||||||||||||||||||||||||| :|| |||   ||||:|||||||||||||||
Db     241 lsllslygihkqkeksrlqggvlvneilnhmkratqipsykklimysahdttvsglqmal 300

Qy     300 DVYNGVLPPYASCHMMELYHDKGGHFVEMYYRNETQNEPYPLTLPGCTHSCPLEKFAELL 359
           |||||:||||||||: ||| :|| :|||||||||||:||||| |||||: |||||:||||:
Db     301 dvyngllppyaschltelyfekgeyfvemyyrnetqhepyplmlpgcspscplerfaelv 360

Qy     360 DPVIPQDWATECMATSSHQGT 380
            |||||||:|||| |:|||||
Db     361 gpvipqdwstecmttnshqgt 381


RESULT   8
AAB74820
ID   AAB74820 standard; Protein; 386 AA.
XX
AC   AAB74820;
XX
DT   14-JUN-2001  (first entry)
```

```
XX
DE   Prostate tumour antigen amino acid sequence for PAP.
XX
KW   Human; prostate tumour antigen; prostate tumour; therapy; diagnosis;
KW   prostate cancer; immunogenic; cytostatic; vaccine.
XX
OS   Homo sapiens.
XX
PN   WO200125272-A2.
XX
PD   12-APR-2001.
XX
PF   04-OCT-2000; 2000WO-US27464.
XX
PR   04-OCT-1999;   99US-0157455.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Xu J,  Skeiky YAW,  Reed SG,  Cheever MA;
XX
DR   WPI; 2001-245062/25.
XX
PT   Prostate specific protein and its encoding polynucleotide, useful for
PT   the treatment and diagnosis of prostate cancer -
XX
PS   Disclosure; Page 270-271; 276pp; English.
XX
CC   The present invention describes an isolated polypeptide (I) comprising
CC   at least an immunogenic portion of a prostate tumour antigen protein or
CC   its variant. (I) have cytostatic activity and can be used in vaccine
CC   production. (I), prostate tumour antigen polynucleotides, an antigen
CC   presenting cell (APC e.g. a dendritic cell) that expresses (I), and a
CC   pharmaceutical composition containing (I) are useful for inhibiting the
CC   development of cancer in a patient. Antibodies specific for prostate
CC   specific proteins and oligonucleotides that hybridise to a
CC   polynucleotide that encodes a prostate specific protein are useful
CC   for detecting the presence or absence of a cancer or monitoring the
CC   progression the progression of a cancer, especially prostate cancer.
CC   AAH02422 to AAH2872, AAB74798 to AAB74821 and AAB74830 are sequences
CC   used in the exemplification of the present invention.
XX
SQ   Sequence   386 AA;


  Query Match             80.7%;  Score 1661.5;  DB 22;  Length 386;
  Best Local Similarity   81.6%;  Pred. No. 2.9e-156;
  Matches  311;  Conservative   28;  Mismatches   41;  Indels    1;  Gaps    1;

Qy       1 MGAVPLPLSPTASLSLGFLLLLSLCLDPG-QAKELKFVTLVFRHGDRGPIETFPTDPITE 59
           | | || |:  |||||||| ||   ||    |||||||||||||||| ||:||||||| |
Db       1 mraaplllaraaslslgflfllffwldrsvlakelkfvtlvfrhgdrspidtfptdpike 60

Qy      60 SSWPQGFGQLTQWGMEQHYELGSYIRKRYGRFLNDTYKHDQIYIRSTDVDRTLMSAMTNL 119
           |||||||||||| |||||||| |||||| :|||::|||:|:|||||||||||||||||||
Db      61 sswpqgfgqltqlgmeqhyelgeyirkryrkflnesykheqvyirstdvdrtlmsamtnl 120

Qy     120 AALFPPEGISIWNPRLLWQPIPVHTVSLSEDRLLYLPFRDCPRFEELKSETLESEEFLKR 179
           ||||||||:||||| |||||||||||| ||||:|||||||:||||:||:||||:|||| ||
Db     121 aalfppegvsiwnpillwqpipvhtvplsedqllylpfrncprfqelesetlkseefqkr 180

Qy     180 LHPYKSFLDTLSSLSGFDDQDLFGIWSKVYDPLFCESVHNFTLPSWATEDAMIKLKELSE 239
           ||||| |: ||  |||   |||||||||||||:|||||||||||||||| | ||:||||
Db     181 lhpykdfiatlgklsglhgqdlfgiwskvydplycesvhnftlpswatedtmtklrelse 240

Qy     240 LSLLSLYGIHKQKEKSRLQGGVLVNEILKNMKLATQPQKYKKLVMYSAHDTTVSGLQMAL 299
           |||||||||||||||||||||||||| :|| |||   ||||:|||||||||||||||||
Db     241 lsllslygihkqkeksrlqggvlvneilnhmkratqipsykklimysahdttvsglqmal 300

Qy     300 DVYNGVLPPYASCHMMELYHDKGGHFVEMYYRNETQNEPYPLTLPGCTHSCPLEKFAELL 359
           |||||:|||||||: ||| :|| :|||||||||||:||||| ||||: |||||:||||:
Db     301 dvyngllppyaschltelyfekgeyfvemyyrnetqhepyplmlpgcspscplerfaelv 360

Qy     360 DPVIPQDWATECMATSSHQGT 380
```

```
           |||||||:|||| |:|||||
Db     361 gpvipqdwstecmttnshqgt 381


RESULT   9
AAW19762
ID   AAW19762 standard; Protein; 515 AA.
XX
AC   AAW19762;
XX
DT   17-SEP-1997  (first entry)
XX
DE   PAP-GM-CSF immunostimulatory fusion protein.
XX
KW   PAP-GM-CSF; granulocyte macrophage colony stimulating factor;
KW   prostatic acid phosphatase; tumour-specific antigen;
KW   immunostimulant; prostate cancer; immunisation; therapy.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Peptide         1..32
FT                   /label= Sig_peptide
FT                   /note= "PAP signal peptide"
FT   Protein         33..515
FT                   /label= Mat_protein
FT                   /note= "PAP-GM-CSF fusion"
FT   Protein         33..386
FT                   /label= PAP
FT   Peptide         387..388
FT                   /label= Linker
FT                   /note= "product of BamHI linker"
FT   Protein         389..515
FT                   /label= GM-CSF
FT   Disulfide-bond  161..372
FT   Disulfide-bond  215..313
FT   Disulfide-bond  347..351
FT   Disulfide-bond  442..484
FT   Disulfide-bond  476..509
FT   Modified-site   94
FT                   /label= Glycosylation
FT                   /note= "potential N-glycosylation site"
FT   Modified-site   220
FT                   /label= Glycosylation
FT                   /note= "potential N-glycosylation site"
FT   Modified-site   333
FT                   /label= Glycosylation
FT                   /note= "potential N-glycosylation site"
FT   Modified-site   415
FT                   /label= Glycosylation
FT                   /note= "potential N-glycosylation site"
FT   Modified-site   425
FT                   /label= Glycosylation
FT                   /note= "potential N-glycosylation site"
XX
PN   WO9724438-A1.
XX
PD   10-JUL-1997.
XX
PF   23-DEC-1996;   96WO-US20241.
XX
PR   28-DEC-1995;   95US-0579823.
XX
PA   (ACTI-) ACTIVATED CELL THERAPY INC.
XX
PI   Laus R,  Ruegg CL,  Wu H;
XX
DR   WPI; 1997-363674/33.
DR   N-PSDB; AAT72721.
XX
PT   Potent APC that activates T-cells to give multivalent cellular
PT   immune response - can also induce a cytotoxic T-cell response in a
PT   vertebrate subject
```

```
XX
PS   Example 1; Fig 2; 45pp; English.
XX
CC   A fusion protein (AAW19762) comprises human prostatic acid
CC   phosphatase (PAP, a tumour-specific antigen) and granulocyte-
CC   macrophage colony stimulating factor (GM-CSF).  It is the
CC   expression product of a nucleic acid molecule (AAT72721) prepd. by
CC   PCR amplification of PAP GM-CSF cDNAs and their fusion via a BamHI
CC   linker.  Fusion expression vectors can be used to transfect
CC   mammalian and insect cells.  The PAP-GM-CSF fusion protein is used
CC   to induce anti-PAP immunity.  PAP serves both as an inducer of
CC   cytotoxic T lymphocytes (CTL) and as a target for prostate cancer
CC   cells when combined with the dendritic cell binding protein GM-CSF
CC   and used to stimulate antigen presenting cells that are then used
CC   to prime CTL.
XX
SQ   Sequence   515 AA;


  Query Match             80.7%;  Score 1661.5;  DB 18;  Length 515;
  Best Local Similarity   81.6%;  Pred. No. 4.6e-156;
  Matches  311;  Conservative   28;  Mismatches   41;  Indels    1;  Gaps    1;

Qy       1 MGAVPLPLSPTASLSLGFLLLLSLCLDPG-QAKELKFVTLVFRHGDRGPIETFPTDPITE 59
           | | || |:  |||||||| ||    ||    ||||||||||||||| ||:||||||| |
Db       1 mraaplllaraaslslgflfllffwldrsvlakelkfvtlvfrhgdrspidtfptdpike 60

Qy      60 SSWPQGFGQLTQWGMEQHYELGSYIRKRYGRFLNDTYKHDQIYIRSTDVDRTLMSAMTNL 119
           |||||||||| |||||||| |||||| :|||::|||:|:|||||||||||||||||||
Db      61 sswpqgfgqltqlgmeqhyelgeyirkryrkflnesykheqvyirstdvdrtlmsamtnl 120

Qy     120 AALFPPEGISIWNPRLLWQPIPVHTVSLSEDRLLYLPFRDCPRFEELKSETLESEEFLKR 179
           ||||||||:||||| |||||||||| ||||:|||||||:||||:|||:||||:|||| ||
Db     121 aalfppegvsiwnpillwqpipvhtvplsedqllylpfrncprfqelesetlkseefqkr 180

Qy     180 LHPYKSFLDTLSSLSGFDDQDLFGIWSKVYDPLFCESVHNFTLPSWATEDAMIKLKELSE 239
           ||||| |: ||  |||   |||||||||||||:|||||||||||||||| | ||:||||
Db     181 lhpykdfiatlgklsglhgqdlfgiwskvydplycesvhnftlpswatedtmtklrelse 240

Qy     240 LSLLSLYGIHKQKEKSRLQGGVLVNEILKNMKLATQPQKYKKLVMYSAHDTTVSGLQMAL 299
           |||||||||||||||||||||||||| :|| |||   ||||:|||||||||||||||||
Db     241 lsllslygihkqkeksrlqggvlvneilnhmkratqipsykklimysahdttvsglqmal 300

Qy     300 DVYNGVLPPYASCHMMELYHDKGGHFVEMYYRNETQNEPYPLTLPGCTHSCPLEKFAELL 359
           |||||:|||||||: ||| :|| :|||||||||||:||||| |||||: |||||:||||:
Db     301 dvyngllppyaschltelyfekgeyfvemyyrnetqhepyplmlpgcspscplerfaelv 360

Qy     360 DPVIPQDWATECMATSSHQGT 380
           |||||||:|||| |:|||||
Db     361 gpvipqdwstecmttnshqgt 381


RESULT  10
AAB56451
ID   AAB56451 standard; Protein; 216 AA.
XX
AC   AAB56451;
XX
DT   13-MAR-2001  (first entry)
XX
DE   Human prostate cancer antigen protein sequence SEQ ID NO:1029.
XX
KW   Human; prostate cancer; prostate cancer antigen; detection; diagnosis;
KW   neuroprotective; cytostatic; cardioactive; immunomodulatory; muscular;
KW   vulnerary; gastrointestinal; nephrotropic; antiinfective; gynaecological;
KW   antibacterial; gene therapy; neural; immune; reproductive; renal;
KW   gastrointestinal; pulmonary; cardiovascular; proliferative disorder;
KW   wound; infectious disease.
XX
OS   Homo sapiens.
XX
PN   WO200055174-A1.
XX
PD   21-SEP-2000.
XX
PF   08-MAR-2000; 2000WO-US05988.
XX
PR   12-MAR-1999;   99US-0124270.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
PA   (ROSE/) ROSEN C A.
XX
PI   Rosen CA,  Ruben SM;
XX
DR   WPI; 2000-587513/55.
DR   N-PSDB; AAF15654.
XX
PT   Prostate cancer associated gene sequences, referred to as prostate
PT   cancer antigens, useful for treatment, prevention, and diagnosis of
PT   disorders such as prostate cancer -
XX
PS   Claim 11; Page 1458-1459; 2338pp; English.
XX
CC   AAF15566 to AAF16505 encode the human prostate cancer associated
CC   proteins, called prostate cancer antigens, given in AAB56363 to AAB57302.
CC   The prostate cancer antigens can have neuroprotective, cytostatic,
CC   cardioactive, immunomodulatory, muscular, vulnerary, gastrointestinal,
CC   nephrotropic, antiinfective, gynaecological and antibacterial activities,
CC   and can be used in gene therapy. The prostate cancer antigen
CC   polynucleotides may be used for detection of prostate cancer, chromosome
CC   identification, as chromosome markers, and for numerous other diagnostic
CC   or research purposes. The prostate cancer antigens may be used to treat
CC   disorders such as neural, immune, muscular, reproductive,
CC   gastrointestinal, pulmonary, cardiovascular, renal, and proliferative
CC   disorders, wounds, and infectious diseases. AAF16506 to AAF16514 to
CC   AAB57303 represent sequences used in the exemplification of the present
CC   invention.
XX
SQ   Sequence   216 AA;


  Query Match             45.6%;  Score 939;  DB 21;  Length 216;
  Best Local Similarity   82.0%;  Pred. No. 7.7e-85;
  Matches  173;  Conservative   16;  Mismatches   22;  Indels    0;  Gaps    0;

Qy     170 TLESEEFLKRLHPYKSFLDTLSSLSGFDDQDLFGIWSKVYDPLFCESVHNFTLPSWATED 229
           ||:|||| ||||||| |: ||  |||   |||||||||||||:|||||||||||||||
Db       1 tlkseefqkrlhpykdfiatlgklsglhgqdlfgiwskvydplycesvhnftlpswated 60

Qy     230 AMIKLKELSELSLLSLYGIHKQKEKSRLQGGVLVNEILKNMKLATQPQKYKKLVMYSAHD 289
           | ||:|||||||||||||||||||||||||||||| :|| |||   ||||:||||||
Db      61 tmtklrelselsllslygihkqkeksrlqggvlvneilnhmkratqipsykklimysahd 120

Qy     290 TTVSGLQMALDVYNGVLPPYASCHMMELYHDKGGHFVEMYYRNETQNEPYPLTLPGCTHS 349
           ||||||||||||||:|||||||: ||| :|| :|||||||||||:||||| ||||: |
Db     121 ttvsglqmaldvyngllppyaschltelyfekgeyfvemyyrnetqhepyplmlpgcsps 180

Qy     350 CPLEKFAELLDPVIPQDWATECMATSSHQGT 380
           ||||:|||||: ||||||||:|||| |:|||||
Db     181 cplerfaelvgpvipqdwstecmttnshqgt 211


RESULT  11
AAG00216
ID   AAG00216 standard; Protein; 210 AA.
XX
AC   AAG00216;
XX
DT   06-OCT-2000  (first entry)
XX
DE   Human secreted protein, SEQ ID NO: 4297.
XX
KW   Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
KW   gene therapy; chromosome mapping.
XX
OS   Homo sapiens.
```

```
XX
PN   EP1033401-A2.
XX
PD   06-SEP-2000.
XX
PF   21-FEB-2000; 2000EP-0200610.
XX
PR   26-FEB-1999;   99US-0122487.
XX
PA   (GEST ) GENSET.
XX
PI   Dumas Milne Edwards J,  Duclert A,  Giordano J;
XX
DR   WPI; 2000-500381/45.
DR   N-PSDB; AAC00222.
XX
PT   New nucleic acid that is a 5' expressed sequence tag (5' EST) for
PT   obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for
PT   diagnostic, forensic, gene therapy and chromosome mapping procedures -
XX
PS   Claim 13; SEQ ID 4297; 71pp + CD-ROM; English.
XX
CC   The present sequence is a polypeptide encoded by one of a large number
CC   of 5' ESTs derived from mRNAs encoding secreted proteins. The 5' ESTs
CC   were prepared from total human RNAs or polyA+ RNAs derived from 30
CC   different tissues. EST sequences usually correspond mainly to the 3'
CC   untranslated region (UTR) of the mRNA because they are often obtained
CC   from oligo-dT primed cDNA libraries. Such ESTs are not well suited for
CC   isolating cDNA sequences derived from the 5' ends of mRNAs and even in
CC   those cases where longer cDNA sequences have been obtained, the full 5'
CC   UTR is rarely included. 5' ESTs are derived from mRNAs with intact 5'
CC   ends and can therefore be used to obtain full length cDNAs and genomic
CC   DNAs. 5' ESTs are also used in diagnostic, forensic, gene therapy and
CC   chromosome mapping procedures. They are used to obtain upstream
CC   regulatory sequences and to design expression and secretion vectors.
XX
SQ   Sequence   210 AA;


  Query Match             42.7%;  Score 880.5;  DB 21;  Length 210;
  Best Local Similarity   80.0%;  Pred. No. 4.8e-79;
  Matches  168;  Conservative   14;  Mismatches   27;  Indels    1;  Gaps    1;

Qy       1 MGAVPLPLSPTASLSLGFLLLLSLCLDPG-QAKELKFVTLVFRHGDRGPIETFPTDPITE 59
           | | || |:  |||||||| ||   ||    |||||||||||||||| ||:|||||| |
Db       1 mraaplllaraaslslgflfllffwldrsvlakelkfvtlvfrhgdrspidtfptdpike 60

Qy      60 SSWPQGFGQLTQWGMEQHYELGSYIRKRYGRFLNDTYKHDQIYIRSTDVDRTLMSAMTNL 119
           ||||||||||| |||||||| |||||| :|||::|||:|:|||||||||||||||||
Db      61 sswpqgfgqltqlgmeqhyelgeyirkryrkflnesykheqvyirstdvdrtlmsamtnl 120

Qy     120 AALFPPEGISIWNPRLLWQPIPVHTVSLSEDRLLYLPFRDCPRFEELKSETLESEEFLKR 179
           ||||||||:||||| |||||||||| ||||:|||||||:||||:||:||||:|||| ||
Db     121 aalfppegvsiwnpillwqpipvhtvplsedqllylpfrncprfqelesetlkseefqkr 180

Qy     180 LHPYKSFLDTLSSLSGFDDQDLFGIWSKVY 209
           ||||| |: ||  |||   ||||||||||
Db     181 lhpykdfiatlgklsglhgqdlfgiwskvy 210


RESULT  12
ABG23348
ID   ABG23348 standard; Protein; 426 AA.
XX
AC   ABG23348;
XX
DT   18-FEB-2002  (first entry)
XX
DE   Novel human diagnostic protein #23339.
XX
KW   Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW   food supplement; medical imaging; diagnostic; genetic disorder.
XX
OS   Homo sapiens.
XX
PN   WO200175067-A2.
XX
PD   11-OCT-2001.
XX
PF   30-MAR-2001; 2001WO-US08631.
XX
PR   31-MAR-2000; 2000US-0540217.
PR   23-AUG-2000; 2000US-0649167.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Drmanac RT,  Liu C,  Tang YT;
XX
DR   WPI; 2001-639362/73.
DR   N-PSDB; AAS87535.
XX
PT   New isolated polynucleotide and encoded polypeptides, useful in
PT   diagnostics, forensics, gene mapping, identification of mutations
PT   responsible for genetic disorders or other traits and to assess
PT   biodiversity  -
XX
PS   Claim 20; SEQ ID No 53707; 103pp; English.
XX
CC   The invention relates to isolated polynucleotide (I) and
CC   polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC   polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC   and gene mapping, and in recombinant production of (II). The
CC   polynucleotides are also used in diagnostics as expressed sequence tags
CC   for identifying expressed genes. (I) is useful in gene therapy techniques
CC   to restore normal activity of (II) or to treat disease states involving
CC   (II). (II) is useful for generating antibodies against it, detecting or
CC   quantitating a polypeptide in tissue, as molecular weight markers and as
CC   a food supplement. (II) and its binding partners are useful in medical
CC   imaging of sites expressing (II). (I) and (II) are useful for treating
CC   disorders involving aberrant protein expression or biological activity.
CC   The polypeptide and polynucleotide sequences have applications in
CC   diagnostics, forensics, gene mapping, identification of mutations
CC   responsible for genetic disorders or other traits to assess biodiversity
CC   and to produce other types of data and products dependent on DNA and
CC   amino acid sequences. ABG00010-ABG30377 represent novel human
CC   diagnostic amino acid sequences of the invention.
CC   Note: The sequence data for this patent did not appear in the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence   426 AA;


  Query Match             36.2%;  Score 745.5;  DB 22;  Length 426;
  Best Local Similarity   42.1%;  Pred. No. 3.8e-65;
  Matches  159;  Conservative   56;  Mismatches  138;  Indels   25;  Gaps    7;

Qy      16 LGF--------LLLLSLCLDPGQAKE--LKFVTLVFRHGDRGPIETFPTDP---ITESSW 62
           |||         |||| | | |    |  | || |||||||| |: ::| ||   :  : |
Db       4 lgfwghpagpllllllvlppralpegplvfvalvfrhgdraplasypmdphkevastlw 63

Qy      63 PQGFGQLTQWGMEQHYELGSYIRKRYGRFLNDTYKHDQIYIRSTDVDRTLMSAMTNLAAL 122
           |:| ||||  |: |  ||| ::| ||  ||:  |: :::|||||| |||| ||  ||| |
Db      64 prglgqlttegvrqqlelgrflrsryeaflspeyrreevyirstdfdrtlesaqanlagl 123

Qy     123 FPPEGISIWNPRLLWQPIPVHTVSLSEDRLLYLPFRDCPRFEELKSETLESEEFLKRLHP 182
           ||        :|   |:||||||| ::||:||  | | |||: ||  |  |: |: : |
Db     124 fpeaapg--spearwrpipvhtvpvaedkllrfpmrscpryhellreateaaeyqealeg 181

Qy     183 YKSFLDTLSSLSGFD--DQDLFGIWSKVYDPLFCESVHNFTLPSWATEDAMIKLKELSEL 240
           :  ||  | : :|     : |   | || | | |:  |   ||:||: | :  | ::| |
Db     182 wtgflsrlenftglslvgeplrraw-kvldtlmcqqahglplpawaspdvlrtlaqisal 240

Qy     241 SLLSLYGIHKQKEKSRLQGGVLVNEILKNMKLATQPQKYKKLVMYSAHDTTVSGLQMALD 300
           : :  |  :  ||::| ||:|:| || |     :     |:|||||||:|:  || ||
Db     241 digahvgppraaekaqltggillnailanfsrvqrlglplkmvmysahdstllalqgalg 300
```

```
Qy    301 VYNGVLPPYASC-------HMMELYHDKGGHFVEMYYRNETQNEPYPLTLPGCTHSCPLE 353
          :|:|  ||||:|       |:     | |   | ::|||:: : | ||:||||   |||
Db    301 lydghtppyaaclgfefrkhlgnpakdggnvtvslfyrndsahlplplslpgcpapcplg 360

Qy    354 KFAELLDPVIPQDWATEC 371
          :| :|  |  |      |
Db    361 rfyqltaparppahgvtc 378


RESULT  13
AAG00217
ID   AAG00217 standard; Protein; 178 AA.
XX
AC   AAG00217;
XX
DT   06-OCT-2000  (first entry)
XX
DE   Human secreted protein, SEQ ID NO: 4298.
XX
KW   Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
KW   gene therapy; chromosome mapping.
XX
OS   Homo sapiens.
XX
PN   EP1033401-A2.
XX
PD   06-SEP-2000.
XX
PF   21-FEB-2000; 2000EP-0200610.
XX
PR   26-FEB-1999;   99US-0122487.
XX
PA   (GEST ) GENSET.
XX
PI   Dumas Milne Edwards J,  Duclert A,  Giordano J;
XX
DR   WPI; 2000-500381/45.
DR   N-PSDB; AAC00223.
XX
PT   New nucleic acid that is a 5' expressed sequence tag (5' EST) for
PT   obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for
PT   diagnostic, forensic, gene therapy and chromosome mapping procedures -
XX
PS   Claim 13; SEQ ID 4298; 71pp + CD-ROM; English.
XX
CC   The present sequence is a polypeptide encoded by one of a large number
CC   of 5' ESTs derived from mRNAs encoding secreted proteins. The 5' ESTs
CC   were prepared from total human RNAs or polyA+ RNAs derived from 30
CC   different tissues. EST sequences usually correspond mainly to the 3'
CC   untranslated region (UTR) of the mRNA because they are often obtained
CC   from oligo-dT primed cDNA libraries. Such ESTs are not well suited for
CC   isolating cDNA sequences derived from the 5' ends of mRNAs and even in
CC   those cases where longer cDNA sequences have been obtained, the full 5'
CC   UTR is rarely included. 5' ESTs are derived from mRNAs with intact 5'
CC   ends and can therefore be used to obtain full length cDNAs and genomic
CC   DNAs. 5' ESTs are also used in diagnostic, forensic, gene therapy and
CC   chromosome mapping procedures. They are used to obtain upstream
CC   regulatory sequences and to design expression and secretion vectors.
XX
SQ   Sequence   178 AA;


  Query Match             33.5%;  Score 689.5;  DB 21;  Length 178;
  Best Local Similarity   66.2%;  Pred. No. 3.6e-60;
  Matches  139;  Conservative   13;  Mismatches   25;  Indels   33;  Gaps    2;

Qy      1 MGAVPLPLSPTASLSLGFLLLLSLCLDPG-QAKELKFVTLVFRHGDRGPIETFPTDPITE 59
          | | || |:  |||||||| ||   ||    |||||||||:
Db      1 mraaplllaraaslslgfflflffwldrsvlakelkfvtll------------------- 41

Qy     60 SSWPQGFGQLTQWGMEQHYELGSYIRKRYGRFLNDTYKHDQIYIRSTDVDRTLMSAMTNL 119
                  ||||||||| |||||| :|||::||| |:||||||||||||||||||
Db     42 ------------gmeqhyelgeyirkryrkflnesykhxqvyirstdvdrtlmsamtnl 88

Qy    120 AALFPPEGISIWNPRLLWQPIPVHTVSLSEDRLLYLPFRDCPRFEELKSETLESEEFLKR 179
          ||||||||:||||| |||||||||||| ||||:|||||||:||||:||:||||:|||| ||
Db     89 aalfppegvsiwnpillwqpipvhtvplsedqllylpfrncprfqelesetlkseefqkr 148

Qy    180 LHPYKSFLDTLSSLSGFDDQDLFGIWSKVY 209
          |||||  |: ||  |||   ||||||||||
Db    149 lhpykdfiatlgklsglhgqdlfgiwskvy 178


RESULT  14
AAB56447
ID   AAB56447 standard; Protein; 171 AA.
XX
AC   AAB56447;
XX
DT   13-MAR-2001  (first entry)
XX
DE   Human prostate cancer antigen protein sequence SEQ ID NO:1025.
XX
KW   Human; prostate cancer; prostate cancer antigen; detection; diagnosis;
KW   neuroprotective; cytostatic; cardioactive; immunomodulatory; muscular;
KW   vulnerary; gastrointestinal; nephrotropic; antiinfective; gynaecological;
KW   antibacterial; gene therapy; neural; immune; reproductive; renal;
KW   gastrointestinal; pulmonary; cardiovascular; proliferative disorder;
KW   wound; infectious disease.
XX
OS   Homo sapiens.
XX
PN   WO200055174-A1.
XX
PD   21-SEP-2000.
XX
PF   08-MAR-2000; 2000WO-US05988.
XX
PR   12-MAR-1999;   99US-0124270.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
PA   (ROSE/) ROSEN C A.
XX
PI   Rosen CA,  Ruben SM;
XX
DR   WPI; 2000-587513/55.
DR   N-PSDB; AAF15650.
XX
PT   Prostate cancer associated gene sequences, referred to as prostate
PT   cancer antigens, useful for treatment, prevention, and diagnosis of
PT   disorders such as prostate cancer -
XX
PS   Claim 11; Page 1455-1456; 2338pp; English.
XX
CC   AAF15566 to AAF16505 encode the human prostate cancer associated
CC   proteins, called prostate cancer antigens, given in AAB56363 to AAB57302.
CC   The prostate cancer antigens can have neuroprotective, cytostatic,
CC   cardioactive, immunomodulatory, muscular, vulnerary, gastrointestinal,
CC   nephrotropic, antiinfective, gynaecological and antibacterial activities,
CC   and can be used in gene therapy. The prostate cancer antigen
CC   polynucleotides may be used for detection of prostate cancer, chromosome
CC   identification, as chromosome markers, and for numerous other diagnostic
CC   or research purposes. The prostate cancer antigens may be used to treat
CC   disorders such as neural, immune, muscular, reproductive,
CC   gastrointestinal, pulmonary, cardiovascular, renal, and proliferative
CC   disorders, wounds, and infectious diseases. AAF16506 to AAF16514 to
CC   AAB57303 represent sequences used in the exemplification of the present
CC   invention.
XX
SQ   Sequence   171 AA;


  Query Match             32.5%;  Score 669.5;  DB 21;  Length 171;
  Best Local Similarity   79.6%;  Pred. No. 3.3e-58;
  Matches  129;  Conservative   11;  Mismatches   21;  Indels    1;  Gaps    1;
```

```
Qy       1 MGAVPLPLSPTASLSLGFLLLLSLCLDPG-QAKELKFVTLVFRHGDRGPIETFPTDPITE 59
Db      10 mraaplllaraaslslgflfllffwldrsvlakelkfvtlvfrhgdrspidtfptdpike 69

Qy      60 SSWPQGFGQLTQWGMEQHYELGSYIRKRYGRFLNDTYKHDQIYIRSTDVDRTLMSAMTNL 119
Db      70 sswpqgfgqltqlgmeqhyelgeyirkryrkflnesykheqvyirstdvdrtlmsamtnl 129

Qy     120 AALFPPEGISIWNPRLLWQPIPVHTVSLSEDRLLYLPFRDCP 161
Db     130 aalfppegvsiwnpillwqpipvhtvplsedqllyltfqelp 171


RESULT  15
ABB58803
ID   ABB58803 standard; Protein; 438 AA.
XX
AC   ABB58803;
XX
DT   26-MAR-2002  (first entry)
XX
DE   Drosophila melanogaster polypeptide SEQ ID NO 3201.
XX
KW   Drosophila; developmental biology; cell signalling; insecticide;
KW   pharmaceutical.
XX
OS   Drosophila melanogaster.
XX
PN   WO200171042-A2.
XX
PD   27-SEP-2001.
XX
PF   23-MAR-2001; 2001WO-US09231.
XX
PR   23-MAR-2000; 2000US-191637P.
PR   11-JUL-2000; 2000US-0614150.
XX
PA   (PEKE ) PE CORP NY.
XX
PI   Venter JC,  Adams M,  Li PWD,  Myers EW;
XX
DR   WPI; 2001-656860/75.
DR   N-PSDB; ABL02906.
XX
PT   New isolated nucleic acid detection reagent for detecting 1000 or more
PT   genes from Drosophila and for elucidating cell signalling and cell-cell
PT   interactions -
XX
PS   Disclosure; SEQ ID NO 3201; 21pp + Sequence Listing; English.
XX
CC   The invention relates to an isolated nucleic acid detection reagent
CC   capable of detecting 1000 or more genes from Drosophila. The invention is
CC   useful in developmental biology and in elucidating cell signalling and
CC   cell-cell interactions in higher eukaryotes for the development of
CC   insecticides, therapeutics and pharmaceutical drugs. The invention
CC   discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC   sequences (ABL01840-ABL16175) and the encoded proteins
CC   (ABB57737-ABB72072).
CC   The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence   438 AA;


  Query Match             29.9%;  Score 616;  DB 22;  Length 438;
  Best Local Similarity   38.4%;  Pred. No. 2.9e-52;
  Matches  142;  Conservative   60;  Mismatches  112;  Indels   56;  Gaps   15;

Qy      28 PGQAKELKFVTLVFRHGDRGPIETFPTDPITESS-WPQGFGQLTQWGMEQHYELGSYIRK 86
Db      45 pgq---lkfvhviyrhgdrtpvdpyptdpwgdrkfwptgwgdltnlgkqehydlgkwlrn 101

Qy      87 RYGRFLNDTYKHDQIYIRSTDVDRTLMSAMTNLAALFPPEGISIWNPRLLWQPIPVHTVS 146
Db     102 rysnllppiysneniyvqstdvdrtlmsaqsnlaglyepqgediwntdinwqpipihtsp 161

Qy     147 LSEDRLLYLPFRDCPRFE-ELKSETLESEEFLKRL----------------HPYKSFLDT 189
Db     162 eredpilaak-apcpaydyelas--lesspefkaltekhrnlfaylsekggrpvktfida 218

Qy     190 LSSLSGFDDQDLFGIWSKVYDPLFCESVHNFTLPSWATEDAMIKLKELSELSLLS--LYG 247
Db     219 ----------------qylnntlfienlynmtlpkwt---kmvygre--eltyvsnfafa 257

Qy     248 IHKQKEK-SRLQGGVLVNEILKNMKL----ATQPQKYKKLVMYSAHDTTVSGLQMALDVY 302
Db     258 issytrklarlkagpllkdifqrfkekssgslkpd--rsmwvysahdttvasvlnalklf 315

Qy     303 NGVLPPYASCHMMELYHDK-GGHFVEMYYRNETQNEPYPLTLPGCTHSCPLEKFAELLDP 361
Db     316 elhsppytacimmelrvdetntplvsifykntta-eplpldipgcgpscpltklmniyed 374

Qy     362 VIPQDWATEC 371
Db     375 vlpvdwerec 384


Search completed: June  4, 2002, 10:31:19
Job time: 90 sec
```

GenCore version 4.5

OM protein - protein search, using sw model

Run on: June 4, 2002, 10:29:49 ; Search time 14.65 Seconds
(without alignments)
641.901 Million cell updates/sec

Title: US-09-402-845-2
Perfect score: 2060
Sequence: 1 MGAVPLPLSPTASLSLGFLL.....DWATECMATSSHQGTVGALG 385

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1661.5	80.7	386	1	US-08-758-213-1	Sequence 1, Appli
2	1661.5	80.7	386	2	US-08-692-787-48	Sequence 48, Appl
3	1661.5	80.7	386	4	US-09-097-199-48	Sequence 48, Appl
4	1661.5	80.7	515	2	US-09-146-283-2	Sequence 2, Appli
5	1661.5	80.7	515	3	US-08-579-823A-2	Sequence 2, Appli
6	1661.5	80.7	515	4	US-09-344-195-2	Sequence 2, Appli
7	164.5	8.0	398	2	US-08-630-822A-56	Sequence 56, Appl
8	164.5	8.0	398	2	US-09-005-069-56	Sequence 56, Appl
9	121.5	5.9	466	4	US-08-868-435-2	Sequence 2, Appli
10	121.5	5.9	466	4	US-08-744-231-2	Sequence 2, Appli
11	117.5	5.7	153	1	US-08-487-001A-53	Sequence 53, Appl
12	117.5	5.7	153	2	US-08-630-822A-53	Sequence 53, Appl
13	117.5	5.7	153	2	US-09-005-069-53	Sequence 53, Appl
14	114.5	5.6	475	2	US-08-819-825-2	Sequence 2, Appli
15	114.5	5.6	475	4	US-09-163-642-2	Sequence 2, Appli
16	106.5	5.2	172	1	US-08-487-001A-51	Sequence 51, Appl
17	106.5	5.2	172	2	US-08-630-822A-51	Sequence 51, Appl
18	106.5	5.2	172	2	US-09-005-069-51	Sequence 51, Appl
19	104	5.0	463	4	US-08-868-435-29	Sequence 29, Appl
20	104	5.0	463	4	US-08-744-231-29	Sequence 29, Appl
21	102.5	5.0	440	3	US-09-259-214-2	Sequence 2, Appli
22	102.5	5.0	440	4	US-09-318-528-2	Sequence 2, Appli
23	102.5	5.0	440	4	US-09-291-931-2	Sequence 2, Appli
24	102.5	5.0	443	3	US-08-993-359-30	Sequence 30, Appl
25	102.5	5.0	479	1	US-07-923-724-2	Sequence 2, Appli
26	102.5	5.0	479	2	US-08-609-426A-2	Sequence 2, Appli
27	102.5	5.0	479	2	US-08-374-652C-4	Sequence 4, Appli
28	101	4.9	466	4	US-08-868-435-35	Sequence 35, Appl
29	101	4.9	466	4	US-08-744-231-35	Sequence 35, Appl
30	100	4.9	446	1	US-07-627-539G-7	Sequence 7, Appli
31	100	4.9	468	1	US-07-627-539G-2	Sequence 2, Appli
32	99.5	4.8	466	4	US-08-868-435-12	Sequence 12, Appl
33	99.5	4.8	466	4	US-08-744-231-12	Sequence 12, Appl
34	97.5	4.7	331	3	US-08-556-419-25	Sequence 25, Appl
35	97.5	4.7	467	1	US-08-151-574-32	Sequence 32, Appl
36	97.5	4.7	467	1	US-08-146-424-20	Sequence 20, Appl
37	97.5	4.7	467	1	US-08-693-709-2	Sequence 2, Appli
38	97.5	4.7	467	2	US-08-419-448-32	Sequence 32, Appl
39	97.5	4.7	467	2	US-08-819-825-3	Sequence 3, Appli
40	97.5	4.7	467	4	US-09-121-425-2	Sequence 2, Appli
41	97.5	4.7	467	4	US-09-163-642-3	Sequence 3, Appli
42	97.5	4.7	467	4	US-09-233-510-32	Sequence 32, Appl
43	96.5	4.7	442	3	US-08-993-359-26	Sequence 26, Appl
44	96.5	4.7	466	4	US-08-868-435-31	Sequence 31, Appl
45	96.5	4.7	466	4	US-08-744-231-31	Sequence 31, Appl

ALIGNMENTS

```
RESULT   1
US-08-758-213-1
; Sequence 1, Application US/08758213
; Patent No. 5763490
;  GENERAL INFORMATION:
;    APPLICANT:  Lebioda, L.
;    APPLICANT:  Jakob, C.
;    TITLE OF INVENTION:  TREATING PROSTATE CANCER BY INHIBITING PROSTATIC ACID PHOSP
;    NUMBER OF SEQUENCES:  1
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Michael A. Mann, P.A.
;      STREET:  Post Office Box 7908
;      CITY:  Columbia
;      STATE:  South Carolina
;      COUNTRY:  United States
;      ZIP:  29202
;    COMPUTER READABLE FORM:  On attached diskette
;      MEDIUM TYPE:  Diskette 3.5 inch, 1.4 MB storage
;      COMPUTER:  Apple Macintosh TM
;      OPERATING SYSTEM:  Macintosh TM
;      SOFTWARE:  Microsoft Word 5.0 TM
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/758,213
;      FILING DATE:  27-NOV-1996
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/650,659
;      FILING DATE:  20-MAY-1996
;      APPLICATION NUMBER:  08/309,091
;      FILING DATE:  09/20/94
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Mann, Michael A.
;      REGISTRATION NUMBER:  32,825
;      REFERENCE/DOCKET NUMBER:  96-2578
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (803) 254-8472
;      TELEFAX:  (803) 254-3760
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  386 amino acids
;      TYPE:  Amino acid
;      STRANDEDNESS:  No. 5763490 Applicable
;      TOPOLOGY:  Linear
US-08-758-213-1


  Query Match            80.7%;  Score 1661.5;  DB 1;  Length 386;
  Best Local Similarity  81.6%;  Pred. No. 2e-169;
  Matches  311;  Conservative   28;  Mismatches   41;  Indels    1;  Gaps    1;
```

```
Qy      1 MGAVPLPLSPTASLSLGFLLLLSLCLDPG-QAKELKFVTLVFRHGDRGPIETFPTDPITE 59
          | | || |:  |||||||| ||    ||    ||||||||||||||| ||:||||||| |
Db      1 MRAAPLLLARAASLSLGFLFLLFFWLDRSVLAKELKFVTLVFRHGDRSPIDTFPTDPIKE 60

Qy     60 SSWPQGFGQLTQWGMEQHYELGSYIRKRYGRFLNDTYKHDQIYIRSTDVDRTLMSAMTNL 119
          ||||||||||||  |||||||||  ||||||  :|||::|||:|:|||||||||||||||||
Db     61 SSWPQGFGQLTQLGMEQHYELGEYIRKRYRKFLNESYKHEQVYIRSTDVDRTLMSAMTNL 120

Qy    120 AALFPPEGISIWNPRLLWQPIPVHTVSLSEDRLLYLPFRDCPRFEELKSETLESEEFLKR 179
          ||||||||:|||||  ||||||||||  ||||:|||||||:||||:||:||||:|||| ||
Db    121 AALFPPEGVSIWNPILLWQPIPVHTVPLSEDQLLYLPFRNCPRFQELESETLKSEEFQKR 180

Qy    180 LHPYKSFLDTLSSLSGFDDQDLFGIWSKVYDPLFCESVHNFTLPSWATEDAMIKLKELSE 239
          |||||  |: ||  |||   ||||||||||||||:||||||||||||||| | ||:||||
Db    181 LHPYKDFIATLGKLSGLHGQDLFGIWSKVYDPLYCESVHNFTLPSWATEDTMTKLRELSE 240

Qy    240 LSLLSLYGIHKQKEKSRLQGGVLVNEILKNMKLATQPQKYKKLVMYSAHDTTVSGLQMAL 299
          ||||||||||||||||||||||||||||| :|| |||   ||||:|||||||||||||||
Db    241 LSLLSLYGIHKQKEKSRLQGGVLVNEILNHMKRATQIPSYKKLIMYSAHDTTVSGLQMAL 300

Qy    300 DVYNGVLPPYASCHMMELYHDKGGHFVEMYYRNETQNEPYPLTLPGCTHSCPLEKFAELL 359
          |||||:|||||||: ||| :|| :|||||||||||:|||| ||||: |||||:||||:
Db    301 DVYNGLLPPYASCHLTELYFEKGEYFVEMYYRNETQHEPYPLMLPGCSPSCPLERFAELV 360

Qy    360 DPVIPQDWATECMATSSHQGT 380
          |||||||:|||| |:|||||
Db    361 GPVIPQDWSTECMTTNSHQGT 381


RESULT   2
US-08-692-787-48
; Sequence 48, Application US/08692787
; Patent No. 5882864
;  GENERAL INFORMATION:
;    APPLICANT:  An, Gang
;    APPLICANT:  O'Hara, S. Mark
;    APPLICANT:  Ralph, David
;    APPLICANT:  Veltri, Robert
;    TITLE OF INVENTION:  BIOMARKERS AND TARGETS FOR DIAGNOSIS,
;    TITLE OF INVENTION:  PROGNOSIS AND MANAGEMENT OF PROSTATE
;    TITLE OF INVENTION:  DISEASE
;    NUMBER OF SEQUENCES:  82
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Arnold, White & Durkee
;      STREET:  P.O. Box 4433
;      CITY:  Houston
;      STATE:  Texas
;      COUNTRY:  United States of America
;      ZIP:  77210
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/692,787
;      FILING DATE:  Concurrently Herewith
;      CLASSIFICATION:  514
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Corder, Timothy S.
;      REGISTRATION NUMBER:  38,414
;      REFERENCE/DOCKET NUMBER:  UROC:012
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (512) 418-3000
;      TELEFAX:  (512) 474-7577
;  INFORMATION FOR SEQ ID NO:  48:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  386 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-692-787-48
```

```
  Query Match             80.7%;  Score 1661.5;  DB 2;  Length 386;
  Best Local Similarity   81.6%;  Pred. No. 2e-169;
  Matches  311;  Conservative   28;  Mismatches   41;  Indels    1;  Gaps    1;

Qy      1 MGAVPLPLSPTASLSLGFLLLLSLCLDPG-QAKELKFVTLVFRHGDRGPIETFPTDPITE 59
          | | || |:  |||||||| ||    ||    ||||||||||||||| ||:||||||| |
Db      1 MRAAPLLLARAASLSLGFLFLLFFWLDRSVLAKELKFVTLVFRHGDRSPIDTFPTDPIKE 60

Qy     60 SSWPQGFGQLTQWGMEQHYELGSYIRKRYGRFLNDTYKHDQIYIRSTDVDRTLMSAMTNL 119
          ||||||||||||  |||||||||  ||||||  :|||::|||:|:|||||||||||||||||
Db     61 SSWPQGFGQLTQLGMEQHYELGEYIRKRYRKFLNESYKHEQVYIRSTDVDRTLMSAMTNL 120

Qy    120 AALFPPEGISIWNPRLLWQPIPVHTVSLSEDRLLYLPFRDCPRFEELKSETLESEEFLKR 179
          ||||||||:|||||  ||||||||||  ||||:|||||||:||||:||:||||:|||| ||
Db    121 AALFPPEGVSIWNPILLWQPIPVHTVPLSEDQLLYLPFRNCPRFQELESETLKSEEFQKR 180

Qy    180 LHPYKSFLDTLSSLSGFDDQDLFGIWSKVYDPLFCESVHNFTLPSWATEDAMIKLKELSE 239
          |||||  |: ||  |||   ||||||||||||||:||||||||||||||| | ||:||||
Db    181 LHPYKDFIATLGKLSGLHGQDLFGIWSKVYDPLYCESVHNFTLPSWATEDTMTKLRELSE 240

Qy    240 LSLLSLYGIHKQKEKSRLQGGVLVNEILKNMKLATQPQKYKKLVMYSAHDTTVSGLQMAL 299
          ||||||||||||||||||||||||||||| :|| |||   ||||:|||||||||||||||
Db    241 LSLLSLYGIHKQKEKSRLQGGVLVNEILNHMKRATQIPSYKKLIMYSAHDTTVSGLQMAL 300

Qy    300 DVYNGVLPPYASCHMMELYHDKGGHFVEMYYRNETQNEPYPLTLPGCTHSCPLEKFAELL 359
          |||||:|||||||: ||| :|| :|||||||||||:|||| ||||: |||||:||||:
Db    301 DVYNGLLPPYASCHLTELYFEKGEYFVEMYYRNETQHEPYPLMLPGCSPSCPLERFAELV 360

Qy    360 DPVIPQDWATECMATSSHQGT 380
          |||||||:|||| |:|||||
Db    361 GPVIPQDWSTECMTTNSHQGT 381


RESULT   3
US-09-097-199-48
; Sequence 48, Application US/09097199
; Patent No. 6218529
;  GENERAL INFORMATION:
;    APPLICANT:  An, Gang
;    APPLICANT:  O'Hara, S. Mark
;    APPLICANT:  Ralph, David
;    APPLICANT:  Veltri, Robert
;    TITLE OF INVENTION:  BIOMARKERS AND TARGETS FOR DIAGNOSIS,
;    TITLE OF INVENTION:  PROGNOSIS AND MANAGEMENT OF PROSTATE DISEASE
;    NUMBER OF SEQUENCES:  87
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Arnold, White & Durkee
;      STREET:  P.O. Box 4433
;      CITY:  Houston
;      STATE:  Texas
;      COUNTRY:  USA
;      ZIP:  77210
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/097,199
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/692,787
;      FILING DATE:  31-JUL-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Nakashima, Richard A.
;      REGISTRATION NUMBER:  P-42,023
;      REFERENCE/DOCKET NUMBER:  UROC:018
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (512) 418-3000
;      TELEFAX:  (512) 474-7577
;  INFORMATION FOR SEQ ID NO:  48:
```

```
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  386 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-09-097-199-48


  Query Match             80.7%;  Score 1661.5;  DB 4;  Length 386;
  Best Local Similarity   81.6%;  Pred. No. 2e-169;
  Matches  311;  Conservative   28;  Mismatches   41;  Indels    1;  Gaps    1;

Qy       1 MGAVPLPLSPTASLSLGFLLLLSLCLDPG-QAKELKFVTLVFRHGDRGPIETFPTDPITE 59
           | | || |:  |||||||| ||   ||    ||||||||||||||| ||:||||||| |
Db       1 MRAAPLLLARAASLSLGFLFLLFFWLDRSVLAKELKFVTLVFRHGDRSPIDTFPTDPIKE 60

Qy      60 SSWPQGFGQLTQWGMEQHYELGSYIRKRYGRFLNDTYKHDQIYIRSTDVDRTLMSAMTNL 119
           |||||||||||| |||||||| |||||| :|||::|||:|:|||||||||||||||||||
Db      61 SSWPQGFGQLTQLGMEQHYELGEYIRKRYRKFLNESYKHEQVYIRSTDVDRTLMSAMTNL 120

Qy     120 AALFPPEGISIWNPRLLWQPIPVHTVSLSEDRLLYLPFRDCPRFEELKSETLESEEFLKR 179
           ||||||||:||||| |||||||||||| ||||:|||||||:||||:||:||||:|||| ||
Db     121 AALFPPEGVSIWNPILLWQPIPVHTVPLSEDQLLYLPFRNCPRFQELESETLKSEEFQKR 180

Qy     180 LHPYKSFLDTLSSLSGFDDQDLFGIWSKVYDPLFCESVHNFTLPSWATEDAMIKLKELSE 239
           ||||| |: ||  |||   |||||||||||||:||||||||||||||| | ||:||||
Db     181 LHPYKDFIATLGKLSGLHGQDLFGIWSKVYDPLYCESVHNFTLPSWATEDTMTKLRELSE 240

Qy     240 LSLLSLYGIHKQKEKSRLQGGVLVNEILKNMKLATQPQKYKKLVMYSAHDTTVSGLQMAL 299
           ||||||||||||||||||||||||||||| :|| |||   ||||:|||||||||||||||
Db     241 LSLLSLYGIHKQKEKSRLQGGVLVNEILNHMKRATQIPSYKKLIMYSAHDTTVSGLQMAL 300

Qy     300 DVYNGVLPPYASCHMMELYHDKGGHFVEMYYRNETQNEPYPLTLPGCTHSCPLEKFAELL 359
           |||||:|||||||: ||| :|| :|||||||||:||||| |||||: |||||:||||:
Db     301 DVYNGLLPPYASCHLTELYFEKGEYFVEMYYRNETQHEPYPLMLPGCSPSCPLERFAELV 360

Qy     360 DPVIPQDWATECMATSSHQGT 380
           |||||||:|||| |:|||||
Db     361 GPVIPQDWSTECMTTNSHQGT 381


RESULT   4
US-09-146-283-2
; Sequence 2, Application US/09146283
; Patent No. 5976546
;  GENERAL INFORMATION:
;    APPLICANT:  Laus, Reiner
;    APPLICANT:  Ruegg, Curtis L.
;    APPLICANT:  Wu, Hongyu
;    TITLE OF INVENTION:  Immunostimulatory Compositions
;    NUMBER OF SEQUENCES:  10
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Dehlinger & Associates
;      STREET:  350 Cambridge Ave. Suite 250
;      CITY:  Palo Alto
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94306
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/146,283
;      FILING DATE:  03-SEPT-1998
;      CLASSIFICATION:  536
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Judge, Linda R.
;      REGISTRATION NUMBER:  42,702
;      REFERENCE/DOCKET NUMBER:  7636-0010.21
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  650-324-0880
;      TELEFAX:  650-324-0960
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  515 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
;    HYPOTHETICAL:  NO
;    ORIGINAL SOURCE:
;      ORGANISM:  Homo sapiens
;      INDIVIDUAL ISOLATE:  Prostatic acid phosphatase-GM-CSF
;      INDIVIDUAL ISOLATE:  fusion protein; Fig. 1
US-09-146-283-2


  Query Match             80.7%;  Score 1661.5;  DB 2;  Length 515;
  Best Local Similarity   81.6%;  Pred. No. 3.1e-169;
  Matches  311;  Conservative   28;  Mismatches   41;  Indels    1;  Gaps    1;

Qy       1 MGAVPLPLSPTASLSLGFLLLLSLCLDPG-QAKELKFVTLVFRHGDRGPIETFPTDPITE 59
           | | || |:  |||||||| ||   ||  _ ||||||||||||||| ||:||||||| |
Db       1 MRAAPLLLARAASLSLGFLFLLFFWLDRSVLAKELKFVTLVFRHGDRSPIDTFPTDPIKE 60

Qy      60 SSWPQGFGQLTQWGMEQHYELGSYIRKRYGRFLNDTYKHDQIYIRSTDVDRTLMSAMTNL 119
           |||||||||||| |||||||| |||||| :|||::|||:|:|||||||||||||||||||
Db      61 SSWPQGFGQLTQLGMEQHYELGEYIRKRYRKFLNESYKHEQVYIRSTDVDRTLMSAMTNL 120

Qy     120 AALFPPEGISIWNPRLLWQPIPVHTVSLSEDRLLYLPFRDCPRFEELKSETLESEEFLKR 179
           ||||||||:||||| |||||||||||| ||||:|||||||:||||:||:||||:|||| ||
Db     121 AALFPPEGVSIWNPILLWQPIPVHTVPLSEDQLLYLPFRNCPRFQELESETLKSEEFQKR 180

Qy     180 LHPYKSFLDTLSSLSGFDDQDLFGIWSKVYDPLFCESVHNFTLPSWATEDAMIKLKELSE 239
           ||||| |: ||  |||   |||||||||||||:||||||||||||||| | ||:||||
Db     181 LHPYKDFIATLGKLSGLHGQDLFGIWSKVYDPLYCESVHNFTLPSWATEDTMTKLRELSE 240

Qy     240 LSLLSLYGIHKQKEKSRLQGGVLVNEILKNMKLATQPQKYKKLVMYSAHDTTVSGLQMAL 299
           ||||||||||||||||||||||||||||| :|| |||   ||||:|||||||||||||||
Db     241 LSLLSLYGIHKQKEKSRLQGGVLVNEILNHMKRATQIPSYKKLIMYSAHDTTVSGLQMAL 300

Qy     300 DVYNGVLPPYASCHMMELYHDKGGHFVEMYYRNETQNEPYPLTLPGCTHSCPLEKFAELL 359
           |||||:|||||||: ||| :|| :|||||||||:||||| |||||: |||||:||||:
Db     301 DVYNGLLPPYASCHLTELYFEKGEYFVEMYYRNETQHEPYPLMLPGCSPSCPLERFAELV 360

Qy     360 DPVIPQDWATECMATSSHQGT 380
           |||||||:|||| |:|||||
Db     361 GPVIPQDWSTECMTTNSHQGT 381


RESULT   5
US-08-579-823A-2
; Sequence 2, Application US/08579823A
; Patent No. 6080409
;  GENERAL INFORMATION:
;    APPLICANT:  Laus, Reiner
;    APPLICANT:  Ruegg, Curtis L.
;    APPLICANT:  Wu, Hongyu
;    TITLE OF INVENTION:  Immunostimulatory Composition and Method
;    NUMBER OF SEQUENCES:  10
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Dehlinger & Associates
;      STREET:  350 Cambridge Ave. Suite 250
;      CITY:  Palo Alto
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94306
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/579,823A
;      FILING DATE:  03-DEC-1998
```

```
;      CLASSIFICATION: 536
;    ATTORNEY/AGENT INFORMATION:
;      NAME: Judge, Linda R.
;      REGISTRATION NUMBER: 42,702
;      REFERENCE/DOCKET NUMBER: 7636-0010
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE: 650-324-0880
;      TELEFAX: 650-324-0960
;  INFORMATION FOR SEQ ID NO: 2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH: 515 amino acids
;      TYPE: amino acid
;      TOPOLOGY: linear
;    MOLECULE TYPE: protein
;    HYPOTHETICAL: NO
;    ORIGINAL SOURCE:
;      ORGANISM: Homo sapiens
;      INDIVIDUAL ISOLATE: Prostatic acid phosphatase-GM-CSF
;      INDIVIDUAL ISOLATE: fusion protein; Fig. 1
US-08-579-823A-2


  Query Match             80.7%;  Score 1661.5;  DB 3;  Length 515;
  Best Local Similarity   81.6%;  Pred. No. 3.1e-169;
  Matches  311;  Conservative   28;  Mismatches   41;  Indels    1;  Gaps    1;

Qy       1 MGAVPLPLSPTASLSLGFLLLLSLCLDPG-QAKELKFVTLVFRHGDRGPIETFPTDPITE 59
Db       1 MRAAPLLLARAASLSLGFLFLLFFWLDRSVLAKELKFVTLVFRHGDRSPIDTFPTDPIKE 60

Qy      60 SSWPQGFGQLTQWGMEQHYELGSYIRKRYGRFLNDTYKHDQIYIRSTDVDRTLMSAMTNL 119
Db      61 SSWPQGFGQLTQLGMEQHYELGEYIRKRYRKFLNESYKHEQVYIRSTDVDRTLMSAMTNL 120

Qy     120 AALFPPEGISIWNPRLLWQPIPVHTVSLSEDRLLYLPFRDCPRFEELKSETLESEEFLKR 179
Db     121 AALFPPEGVSIWNPILLWQPIPVHTVPLSEDQLLYLPFRNCPRFQELESETLKSEEFQKR 180

Qy     180 LHPYKSFLDTLSSLSGFDDQDLFGIWSKVYDPLFCESVHNFTLPSWATEDAMIKLKELSE 239
Db     181 LHPYKDFIATLGKLSGLHGQDLFGIWSKVYDPLYCESVHNFTLPSWATEDTMTKLRELSE 240

Qy     240 LSLLSLYGIHKQKEKSRLQGGVLVNEILKNMKLATQPQKYKKLVMYSAHDTTVSGLQMAL 299
Db     241 LSLLSLYGIHKQKEKSRLQGGVLVNEILNHMKRATQIPSYKKLIMYSAHDTTVSGLQMAL 300

Qy     300 DVYNGVLPPYASCHMMELYHDKGGHFVEMYYRNETQNEPYPLTLPGCTHSCPLEKFAELL 359
Db     301 DVYNGLLPPYASCHLTELYFEKGEYFVEMYYRNETQHEPYPLMLPGCSPSCPLERFAELV 360

Qy     360 DPVIPQDWATECMATSSHQGT 380
Db     361 GPVIPQDWSTECMTTNSHQGT 381


RESULT   6
US-09-344-195-2
; Sequence 2, Application US/09344195
; Patent No. 6210662
;   GENERAL INFORMATION:
;       APPLICANT: Laus, Reiner
;                  Ruegg, Curtis L.
;                  Wu, Hongyu
;       TITLE OF INVENTION: Immunostimulatory Compositions
;       NUMBER OF SEQUENCES: 10
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: Dehlinger & Associates
;            STREET: 350 Cambridge Ave. Suite 250
;            CITY: Palo Alto
;            STATE: CA
;            COUNTRY: USA
;            ZIP: 94306
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: Floppy disk
;            COMPUTER: IBM PC compatible
;            OPERATING SYSTEM: PC-DOS/MS-DOS
;            SOFTWARE: PatentIn Release #1.0, Version #1.25
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/09/344,195
;            FILING DATE: 24-Jun-1999
;            CLASSIFICATION: <Unknown>
;       PRIOR APPLICATION DATA:
;            APPLICATION NUMBER: US/09/146,283
;            FILING DATE: 03-SEPT-1998
;       ATTORNEY/AGENT INFORMATION:
;            NAME: Judge, Linda R.
;            REGISTRATION NUMBER: 42,702
;            REFERENCE/DOCKET NUMBER: 7636-0010.21
;       TELECOMMUNICATION INFORMATION:
;            TELEPHONE: 650-324-0880
;            TELEFAX: 650-324-0960
;  INFORMATION FOR SEQ ID NO: 2:
;       SEQUENCE CHARACTERISTICS:
;            LENGTH: 515 amino acids
;            TYPE: amino acid
;            TOPOLOGY: linear
;       MOLECULE TYPE: protein
;       HYPOTHETICAL: NO
;       ORIGINAL SOURCE:
;            ORGANISM: Homo sapiens
;            INDIVIDUAL ISOLATE: Prostatic acid phosphatase-GM-CSF
;                                fusion protein; Fig. 1
;       SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-344-195-2


  Query Match             80.7%;  Score 1661.5;  DB 4;  Length 515;
  Best Local Similarity   81.6%;  Pred. No. 3.1e-169;
  Matches  311;  Conservative   28;  Mismatches   41;  Indels    1;  Gaps    1;

Qy       1 MGAVPLPLSPTASLSLGFLLLLSLCLDPG-QAKELKFVTLVFRHGDRGPIETFPTDPITE 59
Db       1 MRAAPLLLARAASLSLGFLFLLFFWLDRSVLAKELKFVTLVFRHGDRSPIDTFPTDPIKE 60

Qy      60 SSWPQGFGQLTQWGMEQHYELGSYIRKRYGRFLNDTYKHDQIYIRSTDVDRTLMSAMTNL 119
Db      61 SSWPQGFGQLTQLGMEQHYELGEYIRKRYRKFLNESYKHEQVYIRSTDVDRTLMSAMTNL 120

Qy     120 AALFPPEGISIWNPRLLWQPIPVHTVSLSEDRLLYLPFRDCPRFEELKSETLESEEFLKR 179
Db     121 AALFPPEGVSIWNPILLWQPIPVHTVPLSEDQLLYLPFRNCPRFQELESETLKSEEFQKR 180

Qy     180 LHPYKSFLDTLSSLSGFDDQDLFGIWSKVYDPLFCESVHNFTLPSWATEDAMIKLKELSE 239
Db     181 LHPYKDFIATLGKLSGLHGQDLFGIWSKVYDPLYCESVHNFTLPSWATEDTMTKLRELSE 240

Qy     240 LSLLSLYGIHKQKEKSRLQGGVLVNEILKNMKLATQPQKYKKLVMYSAHDTTVSGLQMAL 299
Db     241 LSLLSLYGIHKQKEKSRLQGGVLVNEILNHMKRATQIPSYKKLIMYSAHDTTVSGLQMAL 300

Qy     300 DVYNGVLPPYASCHMMELYHDKGGHFVEMYYRNETQNEPYPLTLPGCTHSCPLEKFAELL 359
Db     301 DVYNGLLPPYASCHLTELYFEKGEYFVEMYYRNETQHEPYPLMLPGCSPSCPLERFAELV 360

Qy     360 DPVIPQDWATECMATSSHQGT 380
Db     361 GPVIPQDWSTECMTTNSHQGT 381


RESULT   7
US-08-630-822A-56
; Sequence 56, Application US/08630822A
; Patent No. 5840695
;  GENERAL INFORMATION:
;    APPLICANT:  FRANK, GLENN R.
;    APPLICANT:  HUNTER, SHIRLEY WU
```

```
;    APPLICANT:  WALLENFELS, LYNDA
;    TITLE OF INVENTION:  NOVEL ECTOPARASITE SALIVA PROTEINS
;    TITLE OF INVENTION:  AND APPARATUS TO COLLECT SUCH PROTEINS
;    NUMBER OF SEQUENCES:  107
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Sheridan Ross P.C.
;      STREET:  1700 Lincoln Street, Suite 3500
;      CITY:  Denver
;      STATE:  Colorado
;      COUNTRY:  U.S.A.
;      ZIP:  80203
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/630,822A
;      FILING DATE:  11-APR-1996
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  CONNELL, GARY J.
;      REGISTRATION NUMBER:  32,020
;      REFERENCE/DOCKET NUMBER:  2618-17-C3
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (303) 863-9700
;      TELEFAX:  (303) 863-0223
;  INFORMATION FOR SEQ ID NO:  56:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  398 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-630-822A-56


  Query Match             8.0%;  Score 164.5;  DB 2;  Length 398;
  Best Local Similarity   22.5%;  Pred. No. 4.3e-09;
  Matches   89;  Conservative   52;  Mismatches  152;  Indels  103;  Gaps   19;

Qy      33 ELKFVTLVFRHGDRGPIETFPTDP-ITESSWPQGFGQLTQWGMEQHYELGSYIRKRYGRF 91
           |||||    |     | : :|  | ||    |:    ||  |  :  |:        |:
Db      20 ELKFVFATARGMSHTPCD-YPGGPKITNK--PETSSVLTTAGKNEALEI--------GKL 68

Qy      92 LNDTYKHDQIYIRSTDVDRTLMSAMTNL-----AALFPPEGISIWNPRLLW-QPIPVHTV 145
           |:| ||   : ::  |  :   :  :|        |    |:   :    | | |   |
Db      69 LSDHYK-SNLTVKEWDSSKNYWTLASNTRRSQEGTLIIGSGLEGKSRAAEWSQEIGKKTT 127

Qy     146 --SLSEDRLLYLPF---RDCPRFEELKSETLESEEFLKRLHPYKSFLDTLSSLSGFDDQ- 199
             ||    |  |   ::|| |  :| :    :: ||    | :  | |  :   |
Db     128 FSGFSE----YAKFYSQKECPNF--IKQQLDAVKDLLKSAKEYNTEFDKLKKVYNIDAMK 181

Qy     200 ---------DLFGIWSKVYDPLFCESVHNFTLPSWATEDAMIKLKELSELSLLSLYGIHK 250
                  :   : ||: | :   |: |             |||  ||  : :  |:
Db     182 GPQNVWLAYETLNLQSKL-DQIGLGSMKN------------KLKSFSEKYVWA--GLTS 225

Qy     251 QKEKSRLQGGVLVNEILK---NMKLA----TQPQKYKKLVMYSAHDTTVSGLQM------ 297
                :: || ::|:||    |:|        | | | ::  :|       :|:
Db     226 NDNLRKMSGGRMINDILNDIDNIKKGNGQPNAPGKLKIIIGLTAPRFLAESVQLGSKGTK 285

Qy     298 ---------ALDVYNGVLPPYASC-------------------------HMMELYHDKGG 323
                    | |  |:|  : |                           |::||: ||
Db     286 LNQDQNKLSVLTVPQGILAAFVSAFAPKGTKIENQDLDPSSLYPGQGALHVIELHKDKNQ 345

Qy     324 HFVEMYYRNETQNEPYPLTLPGCTHSCPLEKFAELL 359
             |:: |||  |:|  |: |  |  :|  | |   |
Db     346 WNVKILYRNNDQSELKPMKLAKCGDTCSYETFKSTL 381


RESULT   8
US-09-005-069-56
; Sequence 56, Application US/09005069
; Patent No. 5932470
```

```
;  GENERAL INFORMATION:
;    APPLICANT:  FRANK, GLENN R.
;    APPLICANT:  HUNTER, SHIRLEY WU
;    APPLICANT:  WALLENFELS, LYNDA
;    TITLE OF INVENTION:  NOVEL ECTOPARASITE SALIVA PROTEINS
;    TITLE OF INVENTION:  AND APPARATUS TO COLLECT SUCH PROTEINS
;    NUMBER OF SEQUENCES:  107
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Sheridan Ross P.C.
;      STREET:  1700 Lincoln Street, Suite 3500
;      CITY:  Denver
;      STATE:  Colorado
;      COUNTRY:  U.S.A.
;      ZIP:  80203
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/005,069
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/630,822
;      FILING DATE:  11-APR-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  CONNELL, GARY J.
;      REGISTRATION NUMBER:  32,020
;      REFERENCE/DOCKET NUMBER:  2618-17-C3
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (303) 863-9700
;      TELEFAX:  (303) 863-0223
;  INFORMATION FOR SEQ ID NO:  56:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  398 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-09-005-069-56


  Query Match             8.0%;  Score 164.5;  DB 2;  Length 398;
  Best Local Similarity   22.5%;  Pred. No. 4.3e-09;
  Matches   89;  Conservative   52;  Mismatches  152;  Indels  103;  Gaps   19;

Qy      33 ELKFVTLVFRHGDRGPIETFPTDP-ITESSWPQGFGQLTQWGMEQHYELGSYIRKRYGRF 91
           |||||    |     | : :|  | ||    |:    ||  |  :  |:        |:
Db      20 ELKFVFATARGMSHTPCD-YPGGPKITNK--PETSSVLTTAGKNEALEI--------GKL 68

Qy      92 LNDTYKHDQIYIRSTDVDRTLMSAMTNL-----AALFPPEGISIWNPRLLW-QPIPVHTV 145
           |:| ||   : ::  |  :   :  :|        |    |:   :    | | |   |
Db      69 LSDHYK-SNLTVKEWDSSKNYWTLASNTRRSQEGTLIIGSGLEGKSRAAEWSQEIGKKTT 127

Qy     146 --SLSEDRLLYLPF---RDCPRFEELKSETLESEEFLKRLHPYKSFLDTLSSLSGFDDQ- 199
             ||    |  |   ::|| |  :| :    :: ||    | :  | |  :   |
Db     128 FSGFSE----YAKFYSQKECPNF--IKQQLDAVKDLLKSAKEYNTEFDKLKKVYNIDAMK 181

Qy     200 ---------DLFGIWSKVYDPLFCESVHNFTLPSWATEDAMIKLKELSELSLLSLYGIHK 250
                  :   : ||: | :   |: |             |||  ||  : :  |:
Db     182 GPQNVWLAYETLNLQSKL-DQIGLGSMKN------------KLKSFSEKYVWA--GLTS 225

Qy     251 QKEKSRLQGGVLVNEILK---NMKLA----TQPQKYKKLVMYSAHDTTVSGLQM------ 297
                :: || ::|:||    |:|        | | | ::  :|       :|:
Db     226 NDNLRKMSGGRMINDILNDIDNIKKGNGQPNAPGKLKIIIGLTAPRFLAESVQLGSKGTK 285

Qy     298 ---------ALDVYNGVLPPYASC-------------------------HMMELYHDKGG 323
                    | |  |:|  : |                           |::||: ||
Db     286 LNQDQNKLSVLTVPQGILAAFVSAFAPKGTKIENQDLDPSSLYPGQGALHVIELHKDKNQ 345

Qy     324 HFVEMYYRNETQNEPYPLTLPGCTHSCPLEKFAELL 359
             |:: |||  |:|  |: |  |  :|  | |   |
Db     346 WNVKILYRNNDQSELKPMKLAKCGDTCSYETFKSTL 381
```

```
RESULT   9
US-08-868-435-2
; Sequence 2, Application US/08868435
; Patent No. 6291221
;  GENERAL INFORMATION:
;    APPLICANT:  Van Loon, Adolphus
;    APPLICANT:  Mitchell, David
;    TITLE OF INVENTION:  POLYPEPTIDES WITH PHYTASE ACTIVITY
;    NUMBER OF SEQUENCES:  35
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Hoffmann-La Roche Inc.
;      STREET:  340 Kingsland Street
;      CITY:  Nutley
;      STATE:  New Jersey
;      COUNTRY:  United States of America
;      ZIP:  07110
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/868,435
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/744,231
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Kass, Alan P
;      REGISTRATION NUMBER:  32142
;      REFERENCE/DOCKET NUMBER:  Case Docket 9339
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (201) 235-4205
;      TELEFAX:  (201) 235-2363
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  466 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-868-435-2


  Query Match             5.9%;  Score 121.5;  DB 4;  Length 466;
  Best Local Similarity   20.3%;  Pred. No. 0.00022;
  Matches   87;  Conservative   48;  Mismatches  142;  Indels  151;  Gaps   19;

Qy      26 LDPGQAKELKFVTLVFRHGDRGPIE--------TFPTDPITESSWPQGFG---------- 67
           ||  :   : || :: ||| | |          |      : :::|  :
Db      65 LDVPEDCHITFVQVLARHGARSPTHSKTKAYAATIAAIQKSATAFPGKYAFLQSYNYSLD 124

Qy      68 --QLTQWGMEQHYELGSYIRKRYGRFLNDTYKHDQIYIRSTDVDRTLMSAMTNLAALFPP 125
             :|| :|  |  :||:   :||    |   :|   ::|:||  |   ||   :
Db     125 SEELTPFGRNQLRDLGAQFYERY----NALTRHINPFVRATDASRVHESAEKFV------ 174

Qy     126 EGISI-------WNPRLLWQPIPVHTVSLSEDRLL--YLPFRDCPRFE---------- 164
           ||           ||    || |   |:: |       |    |  ||
Db     175 EGFQTARQDDHHANPH---QPSPRVDVAIPEGSAYNNTLEHSLCTAFESSTVGDDAVANF 231

Qy     165 --------------ELKSETLESEEF--LKRLHPYK--SFLDTLSSLSGFDDQDLFGIWS 206
                         :|    | :::   |  : |::  |  |   :|| | |
Db     232 TAVFAPAIAQRLEADLPGVQLSTDDVVNLMAMCPFETVSLTDDAHTLSPFCDL------- 284

Qy     207 KVYDPLFCESVHNFTLPSWATEDAMIKLKELSELSLLSLYGIHKQKEKSRLQGGVLVNEI 266
                     ||    |   : :         |||   ||          :||    ||:
Db     285 ------------FTATEWTQYNYL--------LSLDKYYGYGGGNPLGPVQGVGWANEL 323

Qy     267 LK-------------NMKLATQPQKYK-KLVMYS--AHDTTVSGLQMALDVYNGVLP--- 307
           :              |  |   |  :     :|:  :||: :  :  || :|||  |
Db     324 MARLTRAPVHDHTCVNNTLDASPATFPLNATLYADFSHDSNLVSIFWALGLYNGTAPLSQ 383
```

```
Qy     308 -----------------PYAS---CHMMELYHDKGGHFVEMYYRNETQNEPYPLTLPGC 346
                           |:|:     ||:   :|     |  |    :  ||
Db     384 TSVESVSQTDGYAAAWTVPFAARAYVEMMQCRAEK-----EPLVRVLVNDRVMPL----- 433

Qy     347 THSCPLEK 354
             | || :|
Db     434 -HGCPTDK 440


RESULT  10
US-08-744-231-2
; Sequence 2, Application US/08744231
; Patent No. 6358722
;  GENERAL INFORMATION:
;    APPLICANT:  Van Loon, Adolphus
;    APPLICANT:  Mitchell, David
;    TITLE OF INVENTION:  POLYPEPTIDES WITH PHYTASE ACTIVITY
;    NUMBER OF SEQUENCES:  35
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Hoffmann-La Roche Inc.
;      STREET:  340 Kingsland Street
;      CITY:  Nutley
;      STATE:  New Jersey
;      COUNTRY:  United States of America
;      ZIP:  07110
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/744,231
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/424,757
;      FILING DATE:  18-APR-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Kass, Alan P
;      REGISTRATION NUMBER:  32142
;      REFERENCE/DOCKET NUMBER:  Case Docket 9339
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (201) 235-4205
;      TELEFAX:  (201) 235-2363
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  466 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-744-231-2


  Query Match             5.9%;  Score 121.5;  DB 4;  Length 466;
  Best Local Similarity   20.3%;  Pred. No. 0.00022;
  Matches   87;  Conservative   48;  Mismatches  142;  Indels  151;  Gaps   19;

Qy      26 LDPGQAKELKFVTLVFRHGDRGPIE--------TFPTDPITESSWPQGFG---------- 67
           ||  :   : || :: ||| | |          |      : :::|  :
Db      65 LDVPEDCHITFVQVLARHGARSPTHSKTKAYAATIAAIQKSATAFPGKYAFLQSYNYSLD 124

Qy      68 --QLTQWGMEQHYELGSYIRKRYGRFLNDTYKHDQIYIRSTDVDRTLMSAMTNLAALFPP 125
             :|| :|  |  :||:   :||    |   :|   ::|:||  |   ||   :
Db     125 SEELTPFGRNQLRDLGAQFYERY----NALTRHINPFVRATDASRVHESAEKFV------ 174

Qy     126 EGISI-------WNPRLLWQPIPVHTVSLSEDRLL--YLPFRDCPRFE---------- 164
           ||           ||    || |   |:: |       |    |  ||
Db     175 EGFQTARQDDHHANPH---QPSPRVDVAIPEGSAYNNTLEHSLCTAFESSTVGDDAVANF 231

Qy     165 --------------ELKSETLESEEF--LKRLHPYK--SFLDTLSSLSGFDDQDLFGIWS 206
                         :|    | :::   |  : |::  |  |   :|| | |
Db     232 TAVFAPAIAQRLEADLPGVQLSTDDVVNLMAMCPFETVSLTDDAHTLSPFCDL------- 284
```

```
Qy     207 KVYDPLFCESVHNFTLPSWATEDAMIKLKELSELSLLSLYGIHKQKEKSRLQGGVLVNEI 266
                     ||   |   : :        |||   ||         :||    ||:
Db     285 ------------FTATEWTQYNYL--------LSLDKYYGYGGGNPLGPVQGVGWANEL 323

Qy     267 LK-------------NMKLATQPQKYK-KLVMYS--AHDTTVSGLQMALDVYNGVLP--- 307
           :              |  |   |  :     :|:  :||: :  :  || :|||  |
Db     324 MARLTRAPVHDHTCVNNTLDASPATFPLNATLYADFSHDSNLVSIFWALGLYNGTAPLSQ 383

Qy     308 -----------------PYAS---CHMMELYHDKGGHFVEMYYRNETQNEPYPLTLPGC 346
                           |:|:     ||:   :|     |   |    :   ||
Db     384 TSVESVSQTDGYAAAWTVPFAARAYVEMMQCRAEK-----EPLVRVLVNDRVMPL----- 433

Qy     347 THSCPLEK 354
           | || :|
Db     434 -HGCPTDK 440
```

```
RESULT  11
US-08-487-001A-53
; Sequence 53, Application US/08487001A
; Patent No. 5795862
;  GENERAL INFORMATION:
;    APPLICANT:  FRANK, GLENN R.
;    APPLICANT:  HUNTER, SHIRLEY WU
;    APPLICANT:  WALLENFELS, LYNDA
;    TITLE OF INVENTION:  NOVEL ECTOPARASITE SALIVA
;    TITLE OF INVENTION:  PROTEINS AND APPARATUS TO COLLECT SUCH PROTEINS
;    NUMBER OF SEQUENCES:  54
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Sheridan Ross & McIntosh
;      STREET:  1700 Lincoln Street, Suite 3500
;      CITY:  Denver
;      STATE:  Colorado
;      COUNTRY:  U.S.A.
;      ZIP:  80203
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/487,001A
;      FILING DATE:  07-JUN-1995
;      CLASSIFICATION:  424
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Verser, Carol Talkington
;      REGISTRATION NUMBER:  37,459
;      REFERENCE/DOCKET NUMBER:  2618-17-C2
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (303) 863-9700
;      TELEFAX:  (303) 863-0223
;  INFORMATION FOR SEQ ID NO:  53:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  153 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-487-001A-53
```

```
  Query Match             5.7%;  Score 117.5;  DB 1;  Length 153;
  Best Local Similarity   22.2%;  Pred. No. 0.0001;
  Matches   34;  Conservative   29;  Mismatches   63;  Indels   27;  Gaps    4;

Qy     233 KLKELSELSLLSLYGIHKQKEKSRLQGGVLVNEILKN-------------------MKL 272
           :||  ||  |     :   :   :: || ::|:|| :                     : :
Db       6 ELKSFSEEYLWR--ALTSNENLRKMSGGRMINDILNDIDSIKEERDNRVLLEKQEIKLSM 63

Qy     273 ATQPQKYKKLVM--YSAHDTTVSGLQMALDVYNGVLPPYASCHMMELYHDKGGHFVEMYY 330
           | ||      :  ::   | :    :     : : |   : |::||: |     |:: |
Db      64 LTVPQAILAAFVSAFAPKGTKIENQDLG---PSSLYPGQGALHVIELHKDNNQWSVKVLY 120

Qy     331 RNETQNEPYPLTLPGCTHSCPLEKFAELLDPVI 363
           ||  : |  |: || |   || |   :| :|:|
Db     121 RNNDKMELEPMKLPSCDDKCPCELLNQLYNPMI 153
```

```
RESULT  12
US-08-630-822A-53
; Sequence 53, Application US/08630822A
; Patent No. 5840695
;  GENERAL INFORMATION:
;    APPLICANT:  FRANK, GLENN R.
;    APPLICANT:  HUNTER, SHIRLEY WU
;    APPLICANT:  WALLENFELS, LYNDA
;    TITLE OF INVENTION:  NOVEL ECTOPARASITE SALIVA PROTEINS
;    TITLE OF INVENTION:  AND APPARATUS TO COLLECT SUCH PROTEINS
;    NUMBER OF SEQUENCES:  107
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Sheridan Ross P.C.
;      STREET:  1700 Lincoln Street, Suite 3500
;      CITY:  Denver
;      STATE:  Colorado
;      COUNTRY:  U.S.A.
;      ZIP:  80203
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/630,822A
;      FILING DATE:  11-APR-1996
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  CONNELL, GARY J.
;      REGISTRATION NUMBER:  32,020
;      REFERENCE/DOCKET NUMBER:  2618-17-C3
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (303) 863-9700
;      TELEFAX:  (303) 863-0223
;  INFORMATION FOR SEQ ID NO:  53:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  153 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-630-822A-53
```

```
  Query Match             5.7%;  Score 117.5;  DB 2;  Length 153;
  Best Local Similarity   22.2%;  Pred. No. 0.0001;
  Matches   34;  Conservative   29;  Mismatches   63;  Indels   27;  Gaps    4;

Qy     233 KLKELSELSLLSLYGIHKQKEKSRLQGGVLVNEILKN-------------------MKL 272
           :||  ||  |     :   :   :: || ::|:|| :                     : :
Db       6 ELKSFSEEYLWR--ALTSNENLRKMSGGRMINDILNDIDSIKEERDNRVLLEKQEIKLSM 63

Qy     273 ATQPQKYKKLVM--YSAHDTTVSGLQMALDVYNGVLPPYASCHMMELYHDKGGHFVEMYY 330
           | ||      :  ::   | :    :     : : |   : |::||: |     |:: |
Db      64 LTVPQAILAAFVSAFAPKGTKIENQDLG---PSSLYPGQGALHVIELHKDNNQWSVKVLY 120

Qy     331 RNETQNEPYPLTLPGCTHSCPLEKFAELLDPVI 363
           ||  : |  |: || |   || |   :| :|:|
Db     121 RNNDKMELEPMKLPSCDDKCPCELLNQLYNPMI 153
```

```
RESULT  13
US-09-005-069-53
; Sequence 53, Application US/09005069
; Patent No. 5932470
;  GENERAL INFORMATION:
;    APPLICANT:  FRANK, GLENN R.
;    APPLICANT:  HUNTER, SHIRLEY WU
;    APPLICANT:  WALLENFELS, LYNDA
```

```
;    TITLE OF INVENTION:  NOVEL ECTOPARASITE SALIVA PROTEINS
;    TITLE OF INVENTION:  AND APPARATUS TO COLLECT SUCH PROTEINS
;    NUMBER OF SEQUENCES:  107
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Sheridan Ross P.C.
;      STREET:  1700 Lincoln Street, Suite 3500
;      CITY:  Denver
;      STATE:  Colorado
;      COUNTRY:  U.S.A.
;      ZIP:  80203
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/005,069
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/630,822
;      FILING DATE:  11-APR-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  CONNELL, GARY J.
;      REGISTRATION NUMBER:  32,020
;      REFERENCE/DOCKET NUMBER:  2618-17-C3
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (303) 863-9700
;      TELEFAX:  (303) 863-0223
;  INFORMATION FOR SEQ ID NO:  53:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  153 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-09-005-069-53


  Query Match             5.7%;  Score 117.5;  DB 2;  Length 153;
  Best Local Similarity   22.2%;  Pred. No. 0.0001;
  Matches   34;  Conservative   29;  Mismatches   63;  Indels   27;  Gaps    4;

Qy    233 KLKELSELSLLSLYGIHKQKEKSRLQGGVLVNEILKN-------------------MKL 272
          :||  ||  |     :  :   :: || ::|:|| :                     : :
Db      6 ELKSFSEEYLWR--ALTSNENLRKMSGGRMINDILNDIDSIKEERDNRVLLEKQEIKLSM 63

Qy    273 ATQPQKYKKLVM--YSAHDTTVSGLQMALDVYNGVLPPYASCHMMELYHDKGGHFVEMYY 330
          | ||      :  ::   | :    :     : : |   : |::||: |     |:: |
Db     64 LTVPQAILAAFVSAFAPKGTKIENQDLG---PSSLYPGQGALHVIELHKDNNQWSVKVLY 120

Qy    331 RNETQNEPYPLTLPGCTHSCPLEKFAELLDPVI 363
          ||  : |  |: || |   || |   :| :|:|
Db    121 RNNDKMELEPMKLPSCDDKCPCELLNQLYNPMI 153


RESULT  14
US-08-819-825-2
; Sequence 2, Application US/08819825
; Patent No. 5866118
;  GENERAL INFORMATION:
;    APPLICANT:  Berka, Randy M.
;    APPLICANT:  Ray, Michael W.
;    APPLICANT:  Klotz, Alan V.
;    TITLE OF INVENTION:  Polypeptides Having Phytase Activity
;    TITLE OF INVENTION:  And Nucleic Acids Encoding Same
;    NUMBER OF SEQUENCES:  5
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  No. 5866118o No. 5866118disk of No. 5866118th America, Inc.
;      STREET:  405 Lexington Avenue, Suite 6400
;      CITY:  New York
;      STATE:  New York
;      COUNTRY:  U.S.A.
;      ZIP:  10174-6401
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/819,825
;      FILING DATE:  18-MAR-1997
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Lambiris, Elias J.
;      REGISTRATION NUMBER:  33,728
;      REFERENCE/DOCKET NUMBER:  4758.200-US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  212 867 0123
;      TELEFAX:  212 867 0298
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  475 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
;    FRAGMENT TYPE:  internal
US-08-819-825-2


  Query Match             5.6%;  Score 114.5;  DB 2;  Length 475;
  Best Local Similarity   20.4%;  Pred. No. 0.0013;
  Matches   91;  Conservative   51;  Mismatches  118;  Indels  185;  Gaps   23;

Qy     34 LKFVTLVFRHGDRGPIETFPT--------------DPITE----------SSWPQGFGQ 68
          ::|| :: ||| |     :||               |  ||          ::  |    -
Db     68 VEFVQVLSRHGAR-----YPTAHKSEVYAELLQRIQDTATEFKGDFAFLRDYAYHLGADN 122

Qy     69 LTQWGMEQHYELGSYIRKRY----------------------GRFLNDTYK--------- 97
          ||::| ||  | |     ||                        | |  ::
Db    123 LTRFGEEQMMESGRQFYHRYREQAREIVPFVRAAGSARVIASAEFFNRGFQDAKDRDPRS 182

Qy     98 -HDQ-------IYIRSTDVDRTLMSAMTNLAA------LFPPEGISIWNPRLLWQPIPVH 143
           ||       |    |  : ||   :|  ||          | | : :: ||:| : |  |
Db    183 NKDQAEPVINVIISEETGSNNTL-DGLTCPAAEEAPDPTQPAEFLQVFGPRVL-KKITKH 240

Qy    144 TVSLSEDRLLYLPFRDCPRFEEL-KSETLESEE--FLKRLHPYKSFLDTLSSLSGFDDQD 200
             ::      |   | | | :|   :|: |:   | ::| |
Db    241 MPGVN------LTLEDVPLFMDLCPFDTVGSDPVLFPRQLSP----------------- 276

Qy    201 LFGIWSKVYDPLFCESVHNFTLPSWATEDAMIKLKELSELSLLSLYGIHKQKEKSRLQGG 260
                    ||   | ||   |   |   | :       | :|       ||  |
Db    277 ------------FC---HLFTADDWMAYDYYYTLDKYYSHGGGSAFG------PSR--GV 313

Qy    261 VLVNEILKNM--------------KLATQPQKYK-KLVMYS--AHDTTVSGLQMALDVYN 303
          |||::  |                |   |: :   |:|:  :|| |::|:  |: :||
Db    314 GFVNELIARMTGNLPVKDHTTVNHTLDDNPETFPLDAVLYADFSHDNTMTGIFSAMGLYN 373

Qy    304 GVLP----------------------PYASCHMMELYHDKGGHFVEMYYRNETQNEPY- 339
          |  |                     |:|:   :||  :     |     | ::||:
Db    374 GTKPLSTSKIQPPTGAAADGYAASWTVPFAARAYVELLRCETETSSE--EEEEGEDEPFV 431

Qy    340 ---------PLTLPGCTHSCPLEKF 355
                  ||      | | ::::
Db    432 RVLVNDRVVPL------HGCRVDRW 450


RESULT  15
US-09-163-642-2
; Sequence 2, Application US/09163642
; Patent No. 6221644
;  GENERAL INFORMATION:
;    APPLICANT:  Berka, Randy M.
;    APPLICANT:  Ray, Michael W.
;    APPLICANT:  Klotz, Alan V.
;    TITLE OF INVENTION:  Polypeptides Having Phytase Activity
```

```
;    TITLE OF INVENTION:  And Nucleic Acids Encoding Same
;    NUMBER OF SEQUENCES:  5
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  No. 6221644o No. 6221644disk of No. 6221644th America, Inc.
;      STREET:  405 Lexington Avenue, Suite 6400
;      CITY:  New York
;      STATE:  New York
;      COUNTRY:  U.S.A.
;      ZIP:  10174-6401
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/163,642
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/819,825
;      FILING DATE:  18-MAR-1997
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Lambiris, Elias J.
;      REGISTRATION NUMBER:  33,728
;      REFERENCE/DOCKET NUMBER:  4758.200-US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  212 867 0123
;      TELEFAX:  212 867 0298
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  475 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
;    FRAGMENT TYPE:  internal
US-09-163-642-2

  Query Match             5.6%;  Score 114.5;  DB 4;  Length 475;
  Best Local Similarity   20.4%;  Pred. No. 0.0013;
  Matches   91;  Conservative   51;  Mismatches  118;  Indels  185;  Gaps   23;

Qy      34 LKFVTLVFRHGDRGPIETFPT---------------DPITE----------SSWPQGFGQ 68
           ::|| :: ||| |     :||                 |  ||           ::  |
Db      68 VEFVQVLSRHGAR-----YPTAHKSEVYAELLQRIQDTATEFKGDFAFLRDYAYHLGADN 122

Qy      69 LTQWGMEQHYELGSYIRKRY---------------------GRFLNDTYK--------- 97
             ||::| ||  | |     ||                         | |  ::
Db     123 LTRFGEEQMMESGRQFYHRYREQAREIVPFVRAAGSARVIASAEFFNRGFQDAKDRDPRS 182

Qy      98 -HDQ-------IYIRSTDVDRTLMSAMTNLAA------LFPPEGISIWNPRLLWQPIPVH 143
             ||        |    |  : ||   :|  ||        | | : :: ||:| : |  |
Db     183 NKDQAEPVINVIISEETGSNNTL-DGLTCPAAEEAPDPTQPAEFLQVFGPRVL-KKITKH 240

Qy     144 TVSLSEDRLLYLPFRDCPRFEEL-KSETLESEE--FLKRLHPYKSFLDTLSSLSGFDDQD 200
             ::      |   | | | :|   :|: |:   | ::| |
Db     241 MPGVN------LTLEDVPLFMDLCPFDTVGSDPVLFPRQLSP----------------- 276

Qy     201 LFGIWSKVYDPLFCESVHNFTLPSWATEDAMIKLKELSELSLLSLYGIHKQKEKSRLQGG 260
                   ||   | ||   |   |    | :       | :|       ||  |
Db     277 ------------FC---HLFTADDWMAYDYYYTLDKYYSHGGGSAFG------PSR--GV 313

Qy     261 VLVNEILKNM----------KLATQPQKYK-KLVMYS--AHDTTVSGLQMALDVYN 303
             |||::  |              |   |: :    |:|:  :|| |::|:  |: :||
Db     314 GFVNELIARMTGNLPVKDHTTVNHTLDDNPETFPLDAVLYADFSHDNTMTGIFSAMGLYN 373

Qy     304 GVLP----------------------PYASCHMMELYHDKGGHFVEMYYRNETQNEPY- 339
           |  |                     |:|:   :||   :     |     | ::||:
Db     374 GTKPLSTSKIQPPTGAAADGYAASWTVPFAARAYVELLRCETETSSE--EEEEGEDEPFV 431

Qy     340 ---------PLTLPGCTHSCPLEKF 355
                   ||      | | ::::

Db     432 RVLVNDRVVPL------HGCRVDRW 450


Search completed: June  4, 2002, 10:30:41
Job time: 52 sec
```

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

GenCore version 4.5

OM protein - protein search, using sw model

Run on: June 4, 2002, 10:30:24 ; Search time 31.98 Seconds
(without alignments)
2082.647 Million cell updates/sec

Title: US-09-402-845-2
Perfect score: 2060
Sequence: 1 MGAVPLPLSPTASLSLGFLL.....DWATECMATSSHQGTVGALG 385

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_19:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	2000	97.1	381	11	Q9QXH7	Q9qxh7 mus musculu
2	1656.5	80.4	386	4	Q96KY0	Q96ky0 homo sapien
3	1648.5	80.0	418	4	Q96QM0	Q96qm0 homo sapien
4	1637.5	79.5	386	4	Q96QK9	Q96qk9 homo sapien
5	1026.5	49.8	423	4	Q9BTU7	Q9btu7 homo sapien
6	746.5	36.2	426	4	Q9BZG2	Q9bzg2 homo sapien
7	649	31.5	447	5	Q9U5U0	Q9u5u0 drosophila
8	649	31.5	447	5	Q9TVH9	Q9tvh9 drosophila
9	648	31.5	447	5	Q9U5U4	Q9u5u4 drosophila
10	648	31.5	447	5	Q9U5T9	Q9u5t9 drosophila
11	648	31.5	447	5	Q9U5T7	Q9u5t7 drosophila
12	647.5	31.4	447	5	O97168	O97168 drosophila
13	647	31.4	447	5	Q9U5U6	Q9u5u6 drosophila
14	646	31.4	447	5	Q9U5U1	Q9u5u1 drosophila
15	645	31.3	447	5	Q9U5T8	Q9u5t8 drosophila
16	645	31.3	447	5	Q9U5T5	Q9u5t5 drosophila
17	645	31.3	447	5	Q9U5T2	Q9u5t2 drosophila
18	645	31.3	447	5	Q9U5T1	Q9u5t1 drosophila
19	645	31.3	447	5	Q9TW17	Q9tw17 drosophila
20	645	31.3	447	5	Q9TVX2	Q9tvx2 drosophila
21	644	31.3	447	5	Q9U5V1	Q9u5v1 drosophila
22	644	31.3	447	5	Q9U5V0	Q9u5v0 drosophila
23	644	31.3	447	5	Q9U5U9	Q9u5u9 drosophila
24	644	31.3	447	5	Q9U5U2	Q9u5u2 drosophila
25	644	31.3	447	5	Q9TW53	Q9tw53 drosophila
26	644	31.3	447	5	Q9TW40	Q9tw40 drosophila
27	644	31.3	447	5	Q9TW19	Q9tw19 drosophila
28	643	31.2	447	5	Q9U5U7	Q9u5u7 drosophila
29	643	31.2	447	5	Q9U5U5	Q9u5u5 drosophila
30	643	31.2	447	5	Q9U5T6	Q9u5t6 drosophila
31	643	31.2	447	5	Q9U5T4	Q9u5t4 drosophila
32	643	31.2	447	5	Q9U5T3	Q9u5t3 drosophila
33	641	31.1	447	5	Q9U5U3	Q9u5u3 drosophila
34	640	31.1	447	5	O97186	O97186 drosophila
35	640	31.1	447	5	O97187	O97187 drosophila
36	639	31.0	447	5	Q9TVM1	Q9tvm1 drosophila
37	638	31.0	447	5	Q9U5T0	Q9u5t0 drosophila
38	638	31.0	447	5	Q9U5S8	Q9u5s8 drosophila
39	637	30.9	447	5	Q9U5S9	Q9u5s9 drosophila
40	632	30.7	447	5	Q9U5U8	Q9u5u8 drosophila
41	616	29.9	438	5	Q9VAD0	Q9vad0 drosophila
42	567	27.5	392	5	Q9VVZ9	Q9vvz9 drosophila
43	487	23.6	410	5	Q9VW00	Q9vw00 drosophila
44	487	23.6	440	5	O17373	O17373 caenorhabdi
45	480	23.3	366	5	Q22630	Q22630 caenorhabdi

ALIGNMENTS

RESULT 1
Q9QXH7
ID Q9QXH7 PRELIMINARY; PRT; 381 AA.
AC Q9QXH7;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE PROSTATIC ACID PHOSPHATASE.
GN ACPP.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Crew M.D., Chatta G.S., Borg C.D.;
RT "Sequence and expression of mouse prostatic acid phosphatase.";
RL Submitted (DEC-1999) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AF210243; AAF23171.1; -.
DR HSSP; P20646; 1RPA.
DR MGD; MGI:1928480; Acpp.
DR InterPro; IPR000560; His_acid_phosphtse.
DR Pfam; PF00328; acid_phosphat; 1.
DR PROSITE; PS00616; HIS_ACID_PHOSPHAT_1; 1.
DR PROSITE; PS00778; HIS_ACID_PHOSPHAT_2; 1.
SQ SEQUENCE 381 AA; 43689 MW; 60AD6919D77CB241 CRC64;

Query Match 97.1%; Score 2000; DB 11; Length 381;
Best Local Similarity 98.7%; Pred. No. 2.6e-170;
Matches 374; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

```
Qy      1 MGAVPLPLSPTASLSLGFLLLLSLCLDPGQAKELKFVTLVFRHGDRGPIETFPTDPITES 60
          | ||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 MRAVPLPLSRTASLSLGFLLLLSLCLDPGQAKELKFVTLVFRHGDRGPIETFPTDPITES 60

Qy     61 SWPQGFGQLTQWGMEQHYELGSYIRKRYGRFLNDTYKHDQIYIRSTDVDRTLMSAMTNLA 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 SWPQGFGQLTQWGMEQHYELGSYIRKRYGRFLNDTYKHDQIYIRSTDVDRTLMSAMTNLA 120
```

```
Qy    121 ALFPPEGISIWNPRLLWQPIPVHTVSLSEDRLLYLPFRDCPRFEELKSETLESEEFLKRL 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 ALFPPEGISIWNPRLLWQPIPVHTVSLSEDRLLYLPFRDCPRFEELKSETLESEEFLKRL 180

Qy    181 HPYKSFLDTLSSLSGFDDQDLFGIWSKVYDPLFCESVHNFTLPSWATEDAMIKLKELSEL 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 HPYKSFLDTLSSLSGFDDQDLFGIWSKVYDPLFCESVHNFTLPSWATEDAMIKLKELSEL 240

Qy    241 SLLSLYGIHKQKEKSRLQGGVLVNEILKNMKLATQPQKYKKLVMYSAHDTTVSGLQMALD 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 SLLSLYGIHKQKEKSRLQGGVLVNEILKNMKLATQPQKYKKLVMYSAHDTTVSGLQMALD 300

Qy    301 VYNGVLPPYASCHMMELYHDKGGHFVEMYYRNETQNEPYPLTLPGCTHSCPLEKFAELLD 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||::||
Db    301 VYNGVLPPYASCHMMELYHDKGGHFVEMYYRNETQNEPYPLTLPGCTHSCPLEKFADVLD 360

Qy    361 PVIPQDWATECMATSSHQG 379
          ||| ||||||||||||||
Db    361 PVISQDWATECMATSSHQG 379


RESULT   2
Q96KY0
ID   Q96KY0      PRELIMINARY;      PRT;   386 AA.
AC   Q96KY0;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   ACID PHOSPHATASE, PROSTATE.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=PROSTATE;
RA   Strausberg R.;
RL   Submitted (OCT-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC016344; AAH16344.1; -.
SQ   SEQUENCE   386 AA;  44540 MW;  FE90E10CEBECADEA CRC64;


  Query Match             80.4%;  Score 1656.5;  DB 4;  Length 386;
  Best Local Similarity   81.4%;  Pred. No. 1.3e-139;
  Matches  310;  Conservative   29;  Mismatches   41;  Indels    1;  Gaps    1;

Qy      1 MGAVPLPLSPTASLSLGFLLLLSLCLDPG-QAKELKFVTLVFRHGDRGPIETFPTDPITE 59
          | | || |:  |||||||| ||   ||    ||||||||||||||| ||:||||||| |
Db      1 MRAAPLLLARAASLSLGFLFLLFFWLDRSVLAKELKFVTLVFRHGDRSPIDTFPTDPIKE 60

Qy     60 SSWPQGFGQLTQWGMEQHYELGSYIRKRYGRFLNDTYKHDQIYIRSTDVDRTLMSAMTNL 119
          ||||||||||| |||||||| |||||  :|||::|||:|:|||||||||||||||||||||
Db     61 SSWPQGFGQLTQLGMEQHYELGEYIRKRYRKFLNESYKHEQVYIRSTDVDRTLMSAMTNL 120

Qy    120 AALFPPEGISIWNPRLLWQPIPVHTVSLSEDRLLYLPFRDCPRFEELKSETLESEEFLKR 179
          ||||||||:||||| |||||||||| ||||:|||||||:||||:||:||||:|||| ||
Db    121 AALFPPEGVSIWNPILLWQPIPVHTVPLSEDQLLYLPFRNCPRFQELESETLKSEEFQKR 180

Qy    180 LHPYKSFLDTLSSLSGFDDQDLFGIWSKVYDPLFCESVHNFTLPSWATEDAMIKLKELSE 239
          ||||| |: ||  |||   ||||||||||||||:|||||||||||||||| | ||:||||
Db    181 LHPYKDFIATLGKLSGLHGQDLFGIWSKVYDPLYCESVHNFTLPSWATEDTMTKLRELSE 240

Qy    240 LSLLSLYGIHKQKEKSRLQGGVLVNEILKNMKLATQPQKYKKLVMYSAHDTTVSGLQMAL 299
          ||||||||||||||||||||||||||||||||||| :|| |||   ||||:|||||||||||
Db    241 LSLLSLYGIHKQKEKSRLQGGVLVNEILNHMKRATQIPSYKKLIMYSAHDTTVSGLQMAL 300

Qy    300 DVYNGVLPPYASCHMMELYHDKGGHFVEMYYRNETQNEPYPLTLPGCTHSCPLEKFAELL 359
          |||||:||||||||: ||| :|| :||||:||||||:||||| ||||: |||||:||||:
Db    301 DVYNGLLPPYASCHLTELYFEKGEYFVEMHYRNETQHEPYPLMLPGCSPSCPLERFAELV 360

Qy    360 DPVIPQDWATECMATSSHQGT 380
.         ||||||||:|||| |:|||||
Db    361 GPVIPQDWSTECMTTNSHQGT 381


RESULT   3
Q96QM0
ID   Q96QM0      PRELIMINARY;      PRT;   418 AA.
AC   Q96QM0;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   ACID PHOSPHATASE, PROSTATE.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=PROSTATE;
RA   Strausberg R.;
RL   Submitted (MAY-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC007460; AAH07460.1; -.
SQ   SEQUENCE   418 AA;  48308 MW;  68E10406974E4462 CRC64;


  Query Match             80.0%;  Score 1648.5;  DB 4;  Length 418;
  Best Local Similarity   81.5%;  Pred. No. 7.4e-139;
  Matches  309;  Conservative   27;  Mismatches   42;  Indels    1;  Gaps    1;

Qy      1 MGAVPLPLSPTASLSLGFLLLLSLCLDPG-QAKELKFVTLVFRHGDRGPIETFPTDPITE 59
          | | || |:  |||||||| ||   ||    ||||||||||||||| ||:||||||| |
Db      1 MRAAPLLLARAASLSLGFLFLLFFWLDRSVLAKELKFVTLVFRHGDRSPIDTFPTDPIKE 60

Qy     60 SSWPQGFGQLTQWGMEQHYELGSYIRKRYGRFLNDTYKHDQIYIRSTDVDRTLMSAMTNL 119
          ||||||||||| |||||||| |||||  :|||::|||:|:|||||||||||||||||||||
Db     61 SSWPQGFGQLTQLGMEQHYELGEYIRKRYRKFLNESYKHEQVYIRSTDVDRTLMSAMTNL 120

Qy    120 AALFPPEGISIWNPRLLWQPIPVHTVSLSEDRLLYLPFRDCPRFEELKSETLESEEFLKR 179
          ||||||||:||||| |||||||||| ||||:|||||||:||||:||:||||:|||| ||
Db    121 AALFPPEGVSIWNPILLWQPIPVHTVPLSEDQLLYLPFRNCPRFQELESETLKSEEFQKR 180

Qy    180 LHPYKSFLDTLSSLSGFDDQDLFGIWSKVYDPLFCESVHNFTLPSWATEDAMIKLKELSE 239
          ||||| |: ||  |||   ||||||||||||||:|||||||||||||||| | ||:||||
Db    181 LHPYKDFIATLGKLSGLHGQDLFGIWSKVYDPLYCESVHNFTLPSWATEDTMTKLRELSE 240

Qy    240 LSLLSLYGIHKQKEKSRLQGGVLVNEILKNMKLATQPQKYKKLVMYSAHDTTVSGLQMAL 299
          ||||||||||||||||||||||||||||||||||| :|| |||   ||||:|||||||||||
Db    241 LSLLSLYGIHKQKEKSRLQGGVLVNEILNHMKRATQIPSYKKLIMYSAHDTTVSGLQMAL 300

Qy    300 DVYNGVLPPYASCHMMELYHDKGGHFVEMYYRNETQNEPYPLTLPGCTHSCPLEKFAELL 359
          |||||:||||||||: ||| :|| :||||:||||||:||||| ||||: |||||:||||:
Db    301 DVYNGLLPPYASCHLTELYFEKGEYFVEMYYRNETQHEPYPLMLPGCSPSCPLERFAELA 360

Qy    360 DPVIPQDWATECMATSSHQ 378
           |||||||:|||| |:|||
Db    361 GPVIPQDWSTECMTTNSHQ 379


RESULT   4
Q96QK9
ID   Q96QK9      PRELIMINARY;      PRT;   386 AA.
AC   Q96QK9;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   ACID PHOSPHATASE, PROSTATE.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=PROSTATE;
RA   Strausberg R.;
```

```
RL   Submitted (MAY-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC008493; AAH08493.1; -.
SQ   SEQUENCE   386 AA;  44515 MW;  A4D817CEC1DC1A84 CRC64;


  Query Match           79.5%;  Score 1637.5;  DB 4;  Length 386;
  Best Local Similarity  80.8%;  Pred. No. 6.3e-138;
  Matches  308;  Conservative   29;  Mismatches   43;  Indels    1;  Gaps    1;

Qy       1 MGAVPLPLSPTASLSLGFLLLLSLCLDPG-QAKELKFVTLVFRHGDRGPIETFPTDPITE 59
           | | || |:  |||:|||| ||   ||    |||||||||||||||| ||:||||||| |
Db       1 MRAAPLLLARAASLNLGFLFLLFFWLDRSVLAKELKFVTLVFRHGDRSPIDTFPTDPIKE 60

Qy      60 SSWPQGFGQLTQWGMEQHYELGSYIRKRYGRFLNDTYKHDQIYIRSTDVDRTLMSAMTNL 119
           |||||||||||| |||||||| |||||| :|||::|||:|:|||||||||||||||||||
Db      61 SSWPQGFGQLTQLGMEQHYELGEYIRKRYRKFLNESYKHEQVYIRSTDVDRTLMSAMTNL 120

Qy     120 AALFPPEGISIWNPRLLWQPIPVHTVSLSEDRLLYLPFRDCPRFEELKSETLESEEFLKR 179
           ||| ||||:||||| |||||||||| ||||:|||||||:||||:||:||||:|||| ||
Db     121 AALVPPEGVSIWNPILLWQPIPVHTVPLSEDQLLYLPFRNCPRFQELESETLKSEEFQKR 180

Qy     180 LHPYKSFLDTLSSLSGFDDQDLFGIWSKVYDPLFCESVHNFTLPSWATEDAMIKLKELSE 239
           ||||| |: ||  |||   ||||||||||||||:|||||||||| |||| | ||:||||
Db     181 LHPYKDFIATLGKLSGLHGQDLFGIWSKVYDPLYCESVHNFTLPSRATEDTMTKLRELSE 240

Qy     240 LSLLSLYGIHKQKEKSRLQGGVLVNEILKNMKLATQPQKYKKLVMYSAHDTTVSGLQMAL 299
           ||||||||||||||||||||||||||| :|| |||   ||||:|||||||||||||||||
Db     241 LSLLSLYGIHKQKEKSRLQGGVLVNEILNHMKRATQIPSYKKLIMYSAHDTTVSGLQMAL 300

Qy     300 DVYNGVLPPYASCHMMELYHDKGGHFVEMYYRNETQNEPYPLTLPGCTHSCPLEKFAELL 359
           |||||:|||||||||: ||| :|| :|||||||||||:||||| |||||: |||||:||||:
Db     301 DVYNGLLPPYASCHLTELYFEKGEYFVEMYYRNETQHEPYPLMLPGCSPSCPLERFAELV 360

Qy     360 DPVIPQDWATECMATSSHQGT 380
           |||||||:|||| |:|||||
Db     361 GPVIPQDWSTECMTTNSHQGT 381


RESULT   5
Q9BTU7
ID   Q9BTU7      PRELIMINARY;      PRT;   423 AA.
AC   Q9BTU7;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   ACID PHOSPHATASE 2, LYSOSOMAL.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=KIDNEY, RENAL CELL ADENOCARCINOMA;
RA   Strausberg R.;
RL   Submitted (FEB-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC003160; AAH03160.1; -.
DR   HSSP; P20646; 1RPA.
DR   InterPro; IPR000560; His_acid_phosphtse.
DR   Pfam; PF00328; acid_phosphat; 1.
DR   PROSITE; PS00616; HIS_ACID_PHOSPHAT_1; 1.
DR   PROSITE; PS00778; HIS_ACID_PHOSPHAT_2; 1.
SQ   SEQUENCE   423 AA;  48316 MW;  1A2CBFB40ED3724B CRC64;


  Query Match           49.8%;  Score 1026.5;  DB 4;  Length 423;
  Best Local Similarity  51.6%;  Pred. No. 2.7e-83;
  Matches  190;  Conservative   61;  Mismatches  112;  Indels    5;  Gaps    3;

Qy      14 LSLGFLLLLSLCLDPGQAKELKFVTLVFRHGDRGPIETFPTDPITESSWPQGFGQLTQWG 73
           | ||  |::   : | ||: |:|||||::||||| |::|:| ||  |  |||||||||: |
Db      16 LLLGVNLVV---MPPTQARSLRFVTLLYRHGDRSPVKTYPKDPYQEEEWPQGFGQLTKEG 72

Qy      74 MEQHYELGSYIRKRYGRFLNDTYKHDQIYIRSTDVDRTLMSAMTNLAALFPPEGISIWNP 133
           | ||:|||  :|:||  ||| :|   ::|:|||| |||||||  ||| |||| |:  :||
Db      73 MLQHWELGQALRQRYHGFLNTSYHRQEVYVRSTDFDRTLMSAEANLAGLFPPNGMQRFNP 132

Qy     134 RLLWQPIPVHTVSLSEDRLLYLPFRDCPRFEELKSETLESEEFLKRLHPYKSFLDTLSSL 193
           : ||||||||| ::|||||  |   |||:|:|::|| :: |:          ||| :::
Db     133 NISWQPIPVHTVPITEDRLLKFPLGPCPRYEQLQNETRQTPEYQNESSRNAQFLDMVANE 192

Qy     194 SGFDDQDLFGIWSKVYDPLFCESVHNFTLPSWATEDAMIKLKELSELSLLSLYGIHKQKE 253
           :|  |  |  :|: ||| |||| |   || ||:   | :|  | : |   |:||::| |
Db     193 TGLTDLTLETVWN-VYDTLFCEQTHGLRLPPWASPQTMQRLSRLKDFSFRFLFGIYQQAE 251

Qy     254 KSRLQGGVLVNEILKNMKLATQPQKYKKLVMYSAHDTTVSGLQMALDVYNGVLPPYASCH 313
           |:||||||: :| ||: |     :  ||::|||||||:  |||||||||   ||||||
Db     252 KARLQGGVLLAQIRKNLTLMATTSQLPKLLVYSAHDTTLVALQMALDVYNGEQAPYASCH 311

Qy     314 MMELYHDKGGHF-VEMYYRNETQNEPYPLTLPGCTHSCPLEKFAELLDPVIPQDWATECM 372
           : ||| :  |:| ||||:|||:   |:||:|||| | |||: |  | :||:|:||  ||
Db     312 IFELYQEDSGNFSVEMYFRNESDKAPWPLSLPGCPHRCPLQDFLRLTEPVVPKDWQQECQ 371

Qy     373 ATSSHQGT 380
             |   |
Db     372 LASGPADT 379


RESULT   6
Q9BZG2
ID   Q9BZG2      PRELIMINARY;      PRT;   426 AA.
AC   Q9BZG2;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   ACID PHOSPHATASE.
GN   ACPT.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=21309073; PubMed=11414767;
RA   Yousef G.M., Diamandis M., Jung K., Diamandis E.P.;
RT   "Molecular cloning of a novel human acid phosphatase gene (ACPT) that
RT   is highly expressed in the testis.";
RL   Genomics 74:385-395(2001).
DR   EMBL; AF321918; AAK09393.1; -.
DR   HSSP; P15309; 2HPA.
DR   InterPro; IPR000560; His_acid_phosphtse.
DR   Pfam; PF00328; acid_phosphat; 1.
DR   PROSITE; PS00616; HIS_ACID_PHOSPHAT_1; 1.
DR   PROSITE; PS00778; HIS_ACID_PHOSPHAT_2; UNKNOWN_1.
SQ   SEQUENCE   426 AA;  46089 MW;  BE930398041DB061 CRC64;


  Query Match           36.2%;  Score 746.5;  DB 4;  Length 426;
  Best Local Similarity  42.1%;  Pred. No. 2.7e-58;
  Matches  159;  Conservative   56;  Mismatches  138;  Indels   25;  Gaps    7;

Qy      16 LGF--------LLLLSLCLDPGQAKE--LKFVTLVFRHGDRGPIETFPTDP---ITESSW 62
           |||         |||| | | |    |  | || |||||||| |: ::| ||   :  : |
Db       4 LGFWGHPAGPLLLLLLLVLPPRALPEGPLVFVALVFRHGDRAPLASYPMDPHKEVASTLW 63

Qy      63 PQGFGQLTQWGMEQHYELGSYIRKRYGRFLNDTYKHDQIYIRSTDVDRTLMSAMTNLAAL 122
           |:| ||||  |: |  ||| ::| ||  ||:  |: :::|||||| |||| ||  ||| |
Db      64 PRGLGQLTTEGVRQQLELGRFLRSRYEAFLSPEYRREEVYIRSTDFDRTLESAQANLAGL 123

Qy     123 FPPEGISIWNPRLLWQPIPVHTVSLSEDRLLYLPFRDCPRFEELKSETLESEEFLKRLHP 182
           ||       :|   |:|||||| ::||:||  | | |||: ||  |  |: |: : |
Db     124 FPEAAPG--SPEARWRPIPVHTVPVAEDKLLRFPMRSCPRYHELLREATEAAEYQEALEG 181

Qy     183 YKSFLDTLSSLSGFD--DQDLFGIWSKVYDPLFCESVHNFTLPSWATEDAMIKLKELSEL 240
           :  ||  | : :|     : |   | || | | |:  |   ||:||: | :  | ::| |
Db     182 WTGFLSRLENFTGLSLVGEPLRRAW-KVLDTLMCQQAHGLPLPAWASPDVLRTLAQISAL 240
```

```
Qy    241 SLLSLYGIHKQKEKSRLQGGVLVNEILKNMKLATQPQKYKKLVMYSAHDTTVSGLQMALD 300
          : :  |  :  ||::| ||:|:| || |     :     |:|||||||:|:  || ||
Db    241 DIGAHVGPPRAAEKAQLTGGILLNAILANFSRVQRLGLPLKMVMYSAHDSTLLALQGALG 300

Qy    301 VYNGVLPPYASC-------HMMELYHDKGGHFVEMYYRNETQNEPYPLTLPGCTHSCPLE 353
          :|:|  ||||:|       |:     | |   | ::|||:: : | ||:|||   |||
Db    301 LYDGHTPPYAACLGFEFRKHLGNPAKDGGNVTVSLFYRNDSAHLPLPLSLPGCPAPCPLG 360

Qy    354 KFAELLDPVIPQDWATEC 371
          :| :|  |  |       |
Db    361 RFYQLTAPARPPAHGVSC 378


RESULT   7
Q9U5U0
ID   Q9U5U0      PRELIMINARY;      PRT;   447 AA.
AC   Q9U5U0;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   ACID PHOSPHATASE-1 (EC 3.1.3.2).
GN   ACPH-1.
OS   Drosophila subobscura (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7241;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=J34ST/054;
RX   MEDLINE=99442390; PubMed=10511564;
RA   Navarro-Sabate A., Aguade M., Segarra C.;
RT   "The relationship between allozyme and chromosomal polymorphism
RT   inferred from nucleotide variation at the Acph-1 gene region of
RT   Drosophila subobscura.";
RL   Genetics 153:871-889(1999).
DR   EMBL; AJ389443; CAB59941.1; -.
DR   HSSP; P20646; 1RPA.
DR   FlyBase; FBgn0013885; Dsub\Acph-1.
DR   InterPro; IPR000560; His_acid_phosphtse.
DR   Pfam; PF00328; acid_phosphat; 2.
DR   PROSITE; PS00616; HIS_ACID_PHOSPHAT_1; UNKNOWN_1.
DR   PROSITE; PS00778; HIS_ACID_PHOSPHAT_2; UNKNOWN_1.
KW   Hydrolase.
SQ   SEQUENCE   447 AA;  50998 MW;  FB7CDB4E1EBE5840 CRC64;


  Query Match             31.5%;  Score 649;  DB 5;  Length 447;
  Best Local Similarity   39.6%;  Pred. No. 1.5e-49;
  Matches  147;  Conservative   64;  Mismatches  130;  Indels   30;  Gaps   12;

Qy     28 PGQAKELKFVTLVFRHGDRGPIETFPTDPITESS-WPQGFGQLTQWGMEQHYELGSYIRK 86
          ||   |||| ::|||||| |:: :|||       || |:|||  | |||||||| ::|
Db     54 PG---ELKFAHVIFRHGDRTPVDPYPTDPWNNRKFWPTGWGQLTNLGKEQHYELGKWLRN 110

Qy     87 RYGRFLNDTYKHDQIYIRSTDVDRTLMSAMTNLAALFPPEGISIWNPRLLWQPIPVHTVS 146
          ||   |   | :: |:::||||||||||||| ::|| |: |:|  ||||| |||:|||||
Db    111 RYKSLLGSRYTNEDIFVQSTDVDRTLMSAQSDLAGLYEPQGDDIWNPRIDWQPVPVHTVP 170

Qy    147 LSEDRLLYLPFRDCPRFEELKSETLE-SEEFLKRLHPYKSFLDTLSSLSG------FDDQ 199
            :| :|      || : : :  ||| | ||      |:  |  |:  ||       | |
Db    171 EKDDSILAAK-ASCPAY-DYELATLEASSEFQALYVRYRELLSYLTQNSGRLVKSFIDAQ 228

Qy    200 DLFGIWSKVYDPLFCESVHNFTLPSWATEDAMIKLKELSELSLLSLYGIHKQKEKSRLQG 259
          |        : || | ::| ||| ||  : :   |||: :|  :       :  :||:
Db    229 YL-------NNTLFIEKLYNMTLPVWA--EKVYGKKELTYVSNFAFSIATLTRSMARLKT 279

Qy    260 GVLVNEILK--NMKLATQPQKYKKLVMYSAHDTTVSGLQMALDVYNGVLPPYASCHMMEL 317
          | |: :| :  : ||  | :  : | :|||||||:: :  :| ::    ||||:| |:|:
Db    280 GPLLKDIFERFDKKLNNQLKPDRSLWIYSAHDTTIANVLNSLKLFELHSPPYAACIMLEM 339

Qy    318 -YHDKGGHFVEMYYRNETQNEPYPLTLPGCTHSCPLEKFAELLDPVIPQDWATEC----M 372
           |     | ::|:| |  || || :|||  ||||:   :|  |:| :|  ||    |
Db    340 RVDDSNTPLVSVFYKNTTA-EPLPLDIPGCGLSCPLKTLVKLYQDVLPGNWERECKRSTM 398

Qy    373 ATSSHQGTVGA 383
          :  :  :||
Db    399 MMTYEEANLGA 409


RESULT   8
Q9TVH9
ID   Q9TVH9      PRELIMINARY;      PRT;   447 AA.
AC   Q9TVH9;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   ACID PHOSPHATASE-1 (EC 3.1.3.2).
GN   ACPH-1.
OS   Drosophila subobscura (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7241;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=A1ST/054, AND J8ST/054;
RX   MEDLINE=99442390; PubMed=10511564;
RA   Navarro-Sabate A., Aguade M., Segarra C.;
RT   "The relationship between allozyme and chromosomal polymorphism
RT   inferred from nucleotide variation at the Acph-1 gene region of
RT   Drosophila subobscura.";
RL   Genetics 153:871-889(1999).
DR   EMBL; AJ389447; CAB59945.1; -.
DR   EMBL; AJ389446; CAB59944.1; -.
DR   HSSP; P20646; 1RPA.
DR   FlyBase; FBgn0013885; Dsub\Acph-1.
DR   InterPro; IPR000560; His_acid_phosphtse.
DR   Pfam; PF00328; acid_phosphat; 2.
DR   PROSITE; PS00616; HIS_ACID_PHOSPHAT_1; UNKNOWN_1.
DR   PROSITE; PS00778; HIS_ACID_PHOSPHAT_2; UNKNOWN_1.
KW   Hydrolase.
SQ   SEQUENCE   447 AA;  51040 MW;  E1CCDB4E0FAF5851 CRC64;


  Query Match             31.5%;  Score 649;  DB 5;  Length 447;
  Best Local Similarity   39.6%;  Pred. No. 1.5e-49;
  Matches  147;  Conservative   64;  Mismatches  130;  Indels   30;  Gaps   12;

Qy     28 PGQAKELKFVTLVFRHGDRGPIETFPTDPITESS-WPQGFGQLTQWGMEQHYELGSYIRK 86
          ||   |||| ::|||||| |:: :|||       || |:|||  | |||||||| ::|
Db     54 PG---ELKFAHVIFRHGDRTPVDPYPTDPWNNRKFWPTGWGQLTNLGKEQHYELGKWLRN 110

Qy     87 RYGRFLNDTYKHDQIYIRSTDVDRTLMSAMTNLAALFPPEGISIWNPRLLWQPIPVHTVS 146
          ||   |   | :: |:::||||||||||||| ::|| |: |:|  ||||| |||:|||||
Db    111 RYKSLLGSRYTNEDIFVQSTDVDRTLMSAQSDLAGLYEPQGDDIWNPRIDWQPVPVHTVP 170

Qy    147 LSEDRLLYLPFRDCPRFEELKSETLE-SEEFLKRLHPYKSFLDTLSSLSG------FDDQ 199
            :| :|      || : : :  ||| | ||      |:  |  |:  ||       | |
Db    171 EKDDSILAAK-ASCPAY-DYELATLEASSEFQALYVRYRELLSYLTQNSGRLVKSFIDAQ 228

Qy    200 DLFGIWSKVYDPLFCESVHNFTLPSWATEDAMIKLKELSELSLLSLYGIHKQKEKSRLQG 259
          |        : || | ::| ||| ||  : :   |||: :|  :       :  :||:
Db    229 YL-------NNTLFIEKLYNMTLPVWA--EKVYGKKELTYVSNFAFSIATLTRSMARLKT 279

Qy    260 GVLVNEILK--NMKLATQPQKYKKLVMYSAHDTTVSGLQMALDVYNGVLPPYASCHMMEL 317
          | |: :| :  : ||  | :  : | :|||||||:: :  :| ::    ||||:| |:|:
Db    280 GPLLKDIFERFDKKLNNQLKPDRSLWIYSAHDTTIANVLNSLKLFELHSPPYAACIMLEM 339

Qy    318 -YHDKGGHFVEMYYRNETQNEPYPLTLPGCTHSCPLEKFAELLDPVIPQDWATEC----M 372
           |     | ::|:| |  || || :|||  ||||:   :|  |:| :|  ||    |
Db    340 RVDDSNTPLVSVFYKNTTA-EPLPLDIPGCGLSCPLKTLVKLYQDVLPVNWERECKRSTM 398

Qy    373 ATSSHQGTVGA 383
          :  :  :||
Db    399 MMTYEEANLGA 409
```

```
RESULT   9
Q9U5U4
ID   Q9U5U4      PRELIMINARY;      PRT;   447 AA.
AC   Q9U5U4;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   ACID PHOSPHATASE-1 (EC 3.1.3.2).
GN   ACPH-1.
OS   Drosophila subobscura (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7241;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=J57ST/100;
RX   MEDLINE=99442390; PubMed=10511564;
RA   Navarro-Sabate A., Aguade M., Segarra C.;
RT   "The relationship between allozyme and chromosomal polymorphism
RT   inferred from nucleotide variation at the Acph-1 gene region of
RT   Drosophila subobscura.";
RL   Genetics 153:871-889(1999).
DR   EMBL; AJ389437; CAB59935.1; -.
DR   HSSP; P15309; 2HPA.
DR   FlyBase; FBgn0013885; Dsub\Acph-1.
DR   InterPro; IPR000560; His_acid_phosphtse.
DR   Pfam; PF00328; acid_phosphat; 2.
DR   PROSITE; PS00616; HIS_ACID_PHOSPHAT_1; UNKNOWN_1.
DR   PROSITE; PS00778; HIS_ACID_PHOSPHAT_2; UNKNOWN_1.
KW   Hydrolase.
SQ   SEQUENCE   447 AA;  51049 MW;  D3E545BFB1301859 CRC64;

  Query Match             31.5%;  Score 648;  DB 5;  Length 447;
  Best Local Similarity   39.1%;  Pred. No. 1.8e-49;
  Matches  145;  Conservative   67;  Mismatches  129;  Indels   30;  Gaps   12;

Qy      28 PGQAKELKFVTLVFRHGDRGPIETFPTDPITESS-WPQGFGQLTQWGMEQHYELGSYIRK 86
           ||   ||||  ::|||||| |:: :|||      || |:|||  | |||||| ::|
Db      54 PG---ELKFAHVIFRHGDRTPVDPYPTDPWNNRKFWPTGWGQLTNLGKEQHYELGKWLRN 110

Qy      87 RYGRFLNDTYKHDQIYIRSTDVDRTLMSAMTNLAALFPPEGISIWNPRLLWQPIPVHTVS 146
           ||   |   | :: |:::|||||||||||| ::|| |: |:|  |||||: |||:|||||
Db     111 RYKSLLGSRYTNEDIFVQSTDVDRTLMSAQSDLAGLYEPQGDDIWNPRIDWQPVPVHTVP 170

Qy     147 LSEDRLLYLPFRDCPRFEELKSETLESEEFLKRLH-PYKSFLDTLSSLSG------FDDQ 199
             :| :|      || : : :  |||:    : ||  |:  |  |:  ||       | |
Db     171 EKDDSILAAK-ASCPAY-DYELATLEASSEFQALHVRYRELLSYLTQNSGRLVKSFIDAQ 228

Qy     200 DLFGIWSKVYDPLFCESVHNFTLPSWATEDAMIKLKELSELSLLSLYGIHKQKEKSRLQG 259
           |        : || | ::| ||| ||  : :   :||: :|  :       :  :||:
Db     229 YL-------NNTLFIEKLYNMTLPVWA--EKVYGKEELTYVSNFAFSIATFTRSMARLKT 279

Qy     260 GVLVNEILK--NMKLATQPQKYKKLVMYSAHDTTVSGLQMALDVYNGVLPPYASCHMMEL 317
           | |: :| :  : ||  | :  : | :|||||||:: :  :| ::    ||||:| |:|:
Db     280 GPLLKDIFERFDKKLNNQLKPDRSLWIYSAHDTTIANVLNSLKLFELHSPPYAACIMLEM 339

Qy     318 -YHDKGGHFVEMYYRNETQNEPYPLTLPGCTHSCPLEKFAELLDPVIPQDWATEC----M 372
            |     | ::|:| |  || || :|||  ||||:   :|   |:| :|  ||     |
Db     340 RVDDSNTPLVSVFYKNTTA-EPLPLDIPGCGLSCPLKTLVKLYQDVLPVNWERECKRSTM 398

Qy     373 ATSSHQGTVGA 383
             :  :  :||
Db     399 MMTYEEANLGA 409


RESULT  10
Q9U5T9
ID   Q9U5T9      PRELIMINARY;      PRT;   447 AA.
AC   Q9U5T9;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   ACID PHOSPHATASE-1 (EC 3.1.3.2).
GN   ACPH-1.
OS   Drosophila subobscura (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7241;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=J61ST/054;
RX   MEDLINE=99442390; PubMed=10511564;
RA   Navarro-Sabate A., Aguade M., Segarra C.;
RT   "The relationship between allozyme and chromosomal polymorphism
RT   inferred from nucleotide variation at the Acph-1 gene region of
RT   Drosophila subobscura.";
RL   Genetics 153:871-889(1999).
DR   EMBL; AJ389444; CAB59942.1; -.
DR   HSSP; P20646; 1RPA.
DR   FlyBase; FBgn0013885; Dsub\Acph-1.
DR   InterPro; IPR000560; His_acid_phosphtse.
DR   Pfam; PF00328; acid_phosphat; 2.
DR   PROSITE; PS00616; HIS_ACID_PHOSPHAT_1; UNKNOWN_1.
DR   PROSITE; PS00778; HIS_ACID_PHOSPHAT_2; UNKNOWN_1.
KW   Hydrolase.
SQ   SEQUENCE   447 AA;  51099 MW;  76CB6A4A35A8E4A7 CRC64;

  Query Match             31.5%;  Score 648;  DB 5;  Length 447;
  Best Local Similarity   39.6%;  Pred. No. 1.8e-49;
  Matches  147;  Conservative   64;  Mismatches  130;  Indels   30;  Gaps   12;

Qy      28 PGQAKELKFVTLVFRHGDRGPIETFPTDPITESS-WPQGFGQLTQWGMEQHYELGSYIRK 86
           ||   ||||  ::|||||| |:: :|||      || |:|||  | |||||| ::|
Db      54 PG---ELKFAHVIFRHGDRTPVDPYPTDPWNNRKFWPTGWGQLTNLGKEQHYELGKWLRN 110

Qy      87 RYGRFLNDTYKHDQIYIRSTDVDRTLMSAMTNLAALFPPEGISIWNPRLLWQPIPVHTVS 146
           ||   |   | :: |:::|||||||||||| ::|| |: |:|  |||||: |||:|||||
Db     111 RYKSLLGSRYTNEDIFVQSTDVDRTLMSAQSDLAGLYEPQGDDIWNPRIDWQPVPVHTVP 170

Qy     147 LSEDRLLYLPFRDCPRFEELKSETLE-SEEFLKRLHPYKSFLDTLSSLSG------FDDQ 199
             :| :|      || : : :  ||| | ||      |:  |  |:  ||       | |
Db     171 EKDDSILAAK-ASCPAY-DYELATLEASSEFHSLYVRYRELLSYLTQNSGRLVKSFIDAQ 228

Qy     200 DLFGIWSKVYDPLFCESVHNFTLPSWATEDAMIKLKELSELSLLSLYGIHKQKEKSRLQG 259
           |        : || | ::| ||| ||  : :   |||: :|  :       :  :||:
Db     229 YL-------NNTLFIEKLYNMTLPVWA--EKVYGKKELTYVSNFAFSIATFTRSMARLKT 279

Qy     260 GVLVNEILK--NMKLATQPQKYKKLVMYSAHDTTVSGLQMALDVYNGVLPPYASCHMMEL 317
           | |: :| :  : ||  | :  : | :|||||||:: :  :| ::    ||||:| |:|:
Db     280 GPLLKDIFERFDKKLNNQLKPDRSLWIYSAHDTTIANVLNSLKLFELHSPPYAACIMLEM 339

Qy     318 -YHDKGGHFVEMYYRNETQNEPYPLTLPGCTHSCPLEKFAELLDPVIPQDWATEC----M 372
            |     | ::|:| |  || || :|||  ||||:   :|   |:| :|  ||     |
Db     340 RVDDSNTPLVSVFYKNTTA-EPLPLDIPGCGLSCPLKTLVKLYQDVLPVNWERECKRSTM 398

Qy     373 ATSSHQGTVGA 383
             :  :  :||
Db     399 MMTYEEANLGA 409


RESULT  11
Q9U5T7
ID   Q9U5T7      PRELIMINARY;      PRT;   447 AA.
AC   Q9U5T7;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   ACID PHOSPHATASE-1 (EC 3.1.3.2).
GN   ACPH-1.
OS   Drosophila subobscura (Fruit fly).
```

```
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7241;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=A4ST/054;
RX   MEDLINE=99442390; PubMed=10511564;
RA   Navarro-Sabate A., Aguade M., Segarra C.;
RT   "The relationship between allozyme and chromosomal polymorphism
RT   inferred from nucleotide variation at the Acph-1 gene region of
RT   Drosophila subobscura.";
RL   Genetics 153:871-889(1999).
DR   EMBL; AJ389448; CAB59946.1; -.
DR   HSSP; P20646; 1RPA.
DR   FlyBase; FBgn0013885; Dsub\Acph-1.
DR   InterPro; IPR000560; His_acid_phosphtse.
DR   Pfam; PF00328; acid_phosphat; 2.
DR   PROSITE; PS00616; HIS_ACID_PHOSPHAT_1; UNKNOWN_1.
DR   PROSITE; PS00778; HIS_ACID_PHOSPHAT_2; UNKNOWN_1.
KW   Hydrolase.
SQ   SEQUENCE   447 AA;  51088 MW;  A80B027ACA8D9279 CRC64;


  Query Match           31.5%;  Score 648;  DB 5;  Length 447;
  Best Local Similarity  39.6%;  Pred. No. 1.8e-49;
  Matches  147;  Conservative   64;  Mismatches  130;  Indels   30;  Gaps   12;

Qy      28 PGQAKELKFVTLVFRHGDRGPIETFPTDPITESS-WPQGFGQLTQWGMEQHYELGSYIRK 86
           ||   ||||  ::|||||| |:: :|||      || |:||||  | |||||||  ::|
Db      54 PG---ELKFAHVIFRHGDRTPVDPYPTDPWNNRKFWPTGWGQLTNLGKEQHYELGKWLRN 110

Qy      87 RYGRFLNDTYKHDQIYIRSTDVDRTLMSAMTNLAALFPPEGISIWNPRLLWQPIPVHTVS 146
           ||   |   | :: |:::|||||||||||| ::|| |: |:|  |||||: |||:|||||
Db     111 RYKSLLGSRYTNEDIFVQSTDVDRTLMSAQSDLAGLYEPQGDDIWNPRIDWQPVPVHTVP 170

Qy     147 LSEDRLLYLPFRDCPRFEELKSETLE-SEEFLKRLHPYKSFLDTLSSLSG------FDDQ 199
             :| :|      || : : :  ||| | ||      |:  |  |:  ||       | |
Db     171 EKDDSILAAK-ASCPAY-DYELATLEASSEFQALYVRYRELLSYLTQNSGRLVKSFIDAQ 228

Qy     200 DLFGIWSKVYDPLFCESVHNFTLPSWATEDAMIKLKELSELSLLSLYGIHKQKEKSRLQG 259
           |        : || | ::| ||| ||  : :   |||: :|  :       :  :||:
Db     229 YL-------NNTLFIEKLYNMTLPVWA--EKVYGKKELTYVSNFAFSIATFTRSMARLKT 279

Qy     260 GVLVNEILK--NMKLATQPQKYKKLVMYSAHDTTVSGLQMALDVYNGVLPPYASCHMMEL 317
           | |: :| :  : ||  | :  : | :|||||||:: :  :| ::    ||||:| |:|:
Db     280 GPLLKDIFERFDKKLNNQLKPDRSLWIYSAHDTTIANVLNSLKLFELHSPPYAACIMLEM 339

Qy     318 -YHDKGGHFVEMYYRNETQNEPYPLTLPGCTHSCPLEKFAELLDPVIPQDWATEC----M 372
            |      | ::|:| |  || || :|||  ||||:   :|   |:| :|  ||    |
Db     340 RVDDSNTPLVSVFYKNTTA-EPLPLDIPGCGLSCPLKTLVKLYQDVLPVNWERECKRSTM 398

Qy     373 ATSSHQGTVGA 383
             :  :  :||
Db     399 MMTYEEANLGA 409


RESULT  12
O97168
ID   O97168      PRELIMINARY;      PRT;   447 AA.
AC   O97168;
DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   ACID PHOSPHATASE-1 (EC 3.1.3.2).
GN   ACPH-1.
OS   Drosophila guanche (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7266;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Navarro-Sabate A., Aguade M., Segarra C.;
RT   "The acid phosphatase-1 gene region in the Drosophila species of the
RT   suboscura cluster.";
RL   Submitted (MAR-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; Y18841; CAB38563.1; -.
DR   HSSP; P15309; 2HPA.
DR   FlyBase; FBgn0019218; Dgua\Acph-1.
DR   InterPro; IPR000560; His_acid_phosphtse.
DR   Pfam; PF00328; acid_phosphat; 2.
DR   PROSITE; PS00616; HIS_ACID_PHOSPHAT_1; UNKNOWN_1.
DR   PROSITE; PS00778; HIS_ACID_PHOSPHAT_2; UNKNOWN_1.
KW   Hydrolase.
SQ   SEQUENCE   447 AA;  51073 MW;  1950B73D3C0C8005 CRC64;


  Query Match           31.4%;  Score 647.5;  DB 5;  Length 447;
  Best Local Similarity  37.4%;  Pred. No. 2e-49;
  Matches  154;  Conservative   66;  Mismatches  137;  Indels   55;  Gaps   13;

Qy      12 ASLSLGFLLLLS------------------------LCLDPGQAKELKFVTLVFRHGDR 46
           | | :| | |||                        |   ||   |||| ::||||||
Db      13 ALLVIGVLCLLSFGIGNAVHIPTYGSSEGETRPPPDQLATLPG---ELKFAHVIFRHGDR 69

Qy      47 GPIETFPTDPITESS-WPQGFGQLTQWGMEQHYELGSYIRKRYGRFLNDTYKHDQIYIRS 105
           |:: :||||      || |:|||  | |||||| ::| ||   |   | :: |:::|
Db      70 TPVDPYPTDPWNNRKFWPTGWGQLTNLGKEQHYELGKWLRNRYKSLLGSRYTNEDIFVQS 129

Qy     106 TDVDRTLMSAMTNLAALFPPEGISIWNPRLLWQPIPVHTVSLSEDRLLYLPFRDCPRFEE 165
           |||||||||| ::|| |: |:|  |||||: |||:|||||   :| :|      || : :
Db     130 TDVDRTLMSAQSDLAGLYEPQGDDIWNPRIDWQPVPVHTVPEKDDSILAAK-ASCPAY-D 187

Qy     166 LKSETLE-SEEFLKRLHPYKSFLDTLSSLSG------FDDQDLFGIWSKVYDPLFCESVH 218
           :  ||| | ||      |:  |  |:  ||       | | |        : || | ::
Db     188 YELATLEASSEFQALYVRYRELLSYLTQNSGRLVKSFIDAQYL-------NNTLFIEKLY 240

Qy     219 NFTLPSWATEDAMIKLKELSELSLLSLYGIHKQKEKSRLQGGVLVNEILK--NMKLATQP 276
           | ||| ||  : :   :||: :|  :       :  :||: | |: :| :  : ||  |
Db     241 NMTLPVWA--EKVFGKEELTYVSNFAFSIATFTRSMARLKTGPLLKDIFERFDKKLNNQL 298

Qy     277 QKYKKLVMYSAHDTTVSGLQMALDVYNGVLPPYASCHMMEL-YHDKGGHFVEMYYRNETQ 335
           :  : | :|||||||:: :  :| ::    ||||:| |:|:   |     | ::|:| |
Db     299 KPDRSLWIYSAHDTTIANVLNSLKLFELHSPPYAACIMLEMRVDDSNTPLVSVFYKNTTA 358

Qy     336 NEPYPLTLPGCTHSCPLEKFAELLDPVIPQDWATEC----MATSSHQGTVGA 383
           || || :|||  ||||:   :|   |:| :|  ||    |  :  :  :||
Db     359 -EPLPLDIPGCGLSCPLKTLVKLYQDVLPVNWERECKRSTMMMTYEEANLGA 409


RESULT  13
Q9U5U6
ID   Q9U5U6      PRELIMINARY;      PRT;   447 AA.
AC   Q9U5U6;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   ACID PHOSPHATASE-1 (EC 3.1.3.2).
GN   ACPH-1.
OS   Drosophila subobscura (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7241;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=J51ST/100;
RX   MEDLINE=99442390; PubMed=10511564;
RA   Navarro-Sabate A., Aguade M., Segarra C.;
RT   "The relationship between allozyme and chromosomal polymorphism
RT   inferred from nucleotide variation at the Acph-1 gene region of
RT   Drosophila subobscura.";
RL   Genetics 153:871-889(1999).
DR   EMBL; AJ389432; CAB59930.1; -.
DR   HSSP; P20646; 1RPA.
```

```
DR   FlyBase; FBgn0013885; Dsub\Acph-1.
DR   InterPro; IPR000560; His_acid_phosphtse.
DR   Pfam; PF00328; acid_phosphat; 2.
DR   PROSITE; PS00616; HIS_ACID_PHOSPHAT_1; UNKNOWN_1.
DR   PROSITE; PS00778; HIS_ACID_PHOSPHAT_2; UNKNOWN_1.
KW   Hydrolase.
SQ   SEQUENCE   447 AA;  51049 MW;  10FABA5FAF6EEFF6 CRC64;


  Query Match           31.4%;  Score 647;  DB 5;  Length 447;
  Best Local Similarity   39.6%;  Pred. No. 2.3e-49;
  Matches  147;  Conservative   64;  Mismatches  130;  Indels   30;  Gaps   12;

Qy      28 PGQAKELKFVTLVFRHGDRGPIETFPTDPITESS-WPQGFGQLTQWGMEQHYELGSYIRK 86
           ||   ||||  ::|||||| |:: :||||      || |:|||| | |||||||| ::|
Db      54 PG---ELKFAHVIFRHGDRTPVDPYPTDPWNNRKFWPTGWGQLTNLGKEQHYELGKWLRN 110

Qy      87 RYGRFLNDTYKHDQIYIRSTDVDRTLMSAMTNLAALFPPEGISIWNPRLLWQPIPVHTVS 146
           ||   |   | :: |:::|||||||||||| ::|| |: |:|  |||||: |||:|||||
Db     111 RYKSLLGSRYTNEDIFVQSTDVDRTLMSAQSDLAGLYEPQGDDIWNPRIDWQPVPVHTVP 170

Qy     147 LSEDRLLYLPFRDCPRFEELKSETLE-SEEFLKRLHPYKSFLDTLSSLSG------FDDQ 199
             :| :|      || : : :  ||| | ||      |:  |  |:  ||       | |
Db     171 EKDDSILAAK-ASCPAY-DYELATLEASSEFQALYVRYRELLSYLTQNSGRLVKSFIDAQ 228

Qy     200 DLFGIWSKVYDPLFCESVHNFTLPSWATEDAMIKLKELSELSLLSLYGIHKQKEKSRLQG 259
           |       : || | ::| ||| ||  : :   :||: :|  :       :  |||:
Db     229 YL-------NNTLFIEKLYNMTLPVWA--EKVYGKEELTYVSNFAFSIATFTRSMSRLKT 279

Qy     260 GVLVNEILK--NMKLATQPQKYKKLVMYSAHDTTVSGLQMALDVYNGVLPPYASCHMMEL 317
           | |: :| :  : ||  | :  : | :|||||||:: :  :| ::    ||||:| |:|:
Db     280 GPLLKDIFERFDKKLNNQLKPDRSLWIYSAHDTTIANVLNSLKLFELHSPPYAACIMLEM 339

Qy     318 -YHDKGGHFVEMYYRNETQNEPYPLTLPGCTHSCPLEKFAELLDPVIPQDWATEC----M 372
            |     | ::|:| |  || || :|||  ||||:   :|   |:| :|  ||    |
Db     340 RVDDSNTPLVSVFYKNTTA-EPLPLDIPGCGLSCPLKTLVKLYQDVLPGNWERECKRSTM 398

Qy     373 ATSSHQGTVGA 383
             :  :  :||
Db     399 MMTYEEANLGA 409


RESULT  14
Q9U5U1
ID   Q9U5U1      PRELIMINARY;      PRT;   447 AA.
AC   Q9U5U1;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   ACID PHOSPHATASE-1 (EC 3.1.3.2).
GN   ACPH-1.
OS   Drosophila subobscura (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7241;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=A2ST/100;
RX   MEDLINE=99442390; PubMed=10511564;
RA   Navarro-Sabate A., Aguade M., Segarra C.;
RT   "The relationship between allozyme and chromosomal polymorphism
RT   inferred from nucleotide variation at the Acph-1 gene region of
RT   Drosophila subobscura.";
RL   Genetics 153:871-889(1999).
DR   EMBL; AJ389441; CAB59939.1; -.
DR   HSSP; P20646; 1RPA.
DR   FlyBase; FBgn0013885; Dsub\Acph-1.
DR   InterPro; IPR000560; His_acid_phosphtse.
DR   Pfam; PF00328; acid_phosphat; 2.
DR   PROSITE; PS00616; HIS_ACID_PHOSPHAT_1; UNKNOWN_1.
DR   PROSITE; PS00778; HIS_ACID_PHOSPHAT_2; UNKNOWN_1.
KW   Hydrolase.
SQ   SEQUENCE   447 AA;  51049 MW;  253DF364DC87D375 CRC64;


  Query Match           31.4%;  Score 646;  DB 5;  Length 447;
  Best Local Similarity   38.8%;  Pred. No. 2.8e-49;
  Matches  144;  Conservative   68;  Mismatches  129;  Indels   30;  Gaps   12;

Qy      28 PGQAKELKFVTLVFRHGDRGPIETFPTDPITESS-WPQGFGQLTQWGMEQHYELGSYIRK 86
           ||   ||||  ::|||||| |:: :||||      || |:|||| | |||||||| ::|
Db      54 PG---ELKFAHVIFRHGDRTPVDPYPTDPWNNRKFWPTGWGQLTNLGKEQHYELGKWLRN 110

Qy      87 RYGRFLNDTYKHDQIYIRSTDVDRTLMSAMTNLAALFPPEGISIWNPRLLWQPIPVHTVS 146
           ||   |   | :: |:::|||||||||||| ::|| |: |:|  |||||: |||:|||||
Db     111 RYKSLLGSRYTNEDIFVQSTDVDRTLMSAQSDLAGLYEPQGDDIWNPRIDWQPVPVHTVP 170

Qy     147 LSEDRLLYLPFRDCPRFEELKSETLESEEFLKRLH-PYKSFLDTLSSLSG------FDDQ 199
             :| :|      || : : :  |||:    : ||  |:  :  |:  ||       | |
Db     171 EKDDSILAAK-ASCPAY-DYELATLEASSEFQALHVRYRELISYLTQNSGRLVKSFIDAQ 228

Qy     200 DLFGIWSKVYDPLFCESVHNFTLPSWATEDAMIKLKELSELSLLSLYGIHKQKEKSRLQG 259
           |       : || | ::| ||| ||  : :   :||: :|  :       :  :||:
Db     229 YL-------NNTLFIEKLYNMTLPVWA--EKVYGKEELTYVSNFAFSIATFTRSMARLKT 279

Qy     260 GVLVNEILK--NMKLATQPQKYKKLVMYSAHDTTVSGLQMALDVYNGVLPPYASCHMMEL 317
           | |: :| :  : ||  | :  : | :|||||||:: :  :| ::    ||||:| |:|:
Db     280 GPLLKDIFERFDKKLNNQLKPDRSLWIYSAHDTTIANVLNSLKLFELHSPPYAACIMLEM 339

Qy     318 -YHDKGGHFVEMYYRNETQNEPYPLTLPGCTHSCPLEKFAELLDPVIPQDWATEC----M 372
            |     | ::|:| |  || || :|||  ||||:   :|   |:| :|  ||    |
Db     340 RVDDSNTPLVSVFYKNTTA-EPLPLDIPGCGLSCPLKTLVKLYQDVLPVNWERECKRSTM 398

Qy     373 ATSSHQGTVGA 383
             :  :  :||
Db     399 MMTYEEANLGA 409


RESULT  15
Q9U5T8
ID   Q9U5T8      PRELIMINARY;      PRT;   447 AA.
AC   Q9U5T8;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   ACID PHOSPHATASE-1 (EC 3.1.3.2).
GN   ACPH-1.
OS   Drosophila subobscura (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7241;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=J4ST/054;
RX   MEDLINE=99442390; PubMed=10511564;
RA   Navarro-Sabate A., Aguade M., Segarra C.;
RT   "The relationship between allozyme and chromosomal polymorphism
RT   inferred from nucleotide variation at the Acph-1 gene region of
RT   Drosophila subobscura.";
RL   Genetics 153:871-889(1999).
DR   EMBL; AJ389445; CAB59943.1; -.
DR   HSSP; P20646; 1RPA.
DR   FlyBase; FBgn0013885; Dsub\Acph-1.
DR   InterPro; IPR000560; His_acid_phosphtse.
DR   Pfam; PF00328; acid_phosphat; 2.
DR   PROSITE; PS00616; HIS_ACID_PHOSPHAT_1; UNKNOWN_1.
DR   PROSITE; PS00778; HIS_ACID_PHOSPHAT_2; UNKNOWN_1.
KW   Hydrolase.
SQ   SEQUENCE   447 AA;  50982 MW;  5A320C40C4FE8803 CRC64;


  Query Match           31.3%;  Score 645;  DB 5;  Length 447;
  Best Local Similarity   39.4%;  Pred. No. 3.4e-49;
  Matches  146;  Conservative   65;  Mismatches  130;  Indels   30;  Gaps   12;
```

```
Qy      28 PGQAKELKFVTLVFRHGDRGPIETFPTDPITESS-WPQGFGQLTQWGMEQHYELGSYIRK 86
           ||   ||||  ::|||||| |:: :|||       || |:|||  | |||:||| ::|
Db      54 PG---ELKFAHVIFRHGDRTPVDPYPTDPWNNRKFWPTGWGQLTNLGKEQHFELGKWLRN 110

Qy      87 RYGRFLNDTYKHDQIYIRSTDVDRTLMSAMTNLAALFPPEGISIWNPRLLWQPIPVHTVS 146
           ||   |   | :: |:::|||||||||||| ::|| |: |:|  |||||: |||:|||||
Db     111 RYKSLLGSRYTNEDIFVQSTDVDRTLMSAQSDLAGLYEPQGDDIWNPRIDWQPVPVHTVP 170

Qy     147 LSEDRLLYLPFRDCPRFEELKSETLE-SEEFLKRLHPYKSFLDTLSSLSG------FDDQ 199
             :| :|      || : : :  ||| | ||       |:  |  |:  ||       | |
Db     171 EKDDSILAAK-ASCPAY-DYELATLEASSEFQALYVRYRELLSYLTQNSGRLVKSFIDAQ 228

Qy     200 DLFGIWSKVYDPLFCESVHNFTLPSWATEDAMIKLKELSELSLLSLYGIHKQKEKSRLQG 259
           |        : || | ::| ||| ||  : :   |||: :|  :       :  :||:
Db     229 YL-------NNTLFIEKLYNMTLPVWA--EKVYGKKELTYVSNFAFSIATLTRSMARLKT 279

Qy     260 GVLVNEILK--NMKLATQPQKYKKLVMYSAHDTTVSGLQMALDVYNGVLPPYASCHMMEL 317
           | |: :| :  : ||  | :  : | :|||||||:: :  :| ::    ||||:| |:|:
Db     280 GPLLKDIFERFDKKLNNQLKPDRSLWIYSAHDTTIANVLNSLKLFELHSPPYAACIMLEM 339

Qy     318 -YHDKGGHFVEMYYRNETQNEPYPLTLPGCTHSCPLEKFAELLDPVIPQDWATEC----M 372
           |      | ::|:| |  || || :|||  ||||:   :|   |:| :|  ||    |
Db     340 RVDDSNTPLVSVFYKNTTA-EPLPLDIPGCGLSCPLKTLVKLYQDVLPGNWERECKRSTM 398

Qy     373 ATSSHQGTVGA 383
           :  :  :||
Db     399 MMTYEEANLGA 409


Search completed: June  4, 2002, 10:34:16
Job time: 232 sec
```

GenCore version 4.5

OM protein - protein search, using sw model

Run on: June 4, 2002, 10:29:49 ; Search time 21.01 Seconds
(without alignments)
1760.799 Million cell updates/sec

Title: US-09-402-845-2
Perfect score: 2060
Sequence: 1 MGAVPLPLSPTASLSLGFLL.....DWATECMATSSHQGTVGALG 385

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_71:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1818	88.3	381	2	JH0152	acid phosphatase (
2	1661.5	80.7	386	1	JH0610	acid phosphatase (
3	1022.5	49.6	423	1	S06167	acid phosphatase (
4	1020.5	49.5	421	2	S14742	acid phosphatase (
5	1019.5	49.5	423	2	A33395	acid phosphatase (
6	617	30.0	438	2	S64682	acid phosphatase (
7	542.5	26.3	411	2	D88504	protein B0361.7 [i
8	487	23.6	440	2	T32457	hypothetical prote
9	480	23.3	366	2	T25060	hypothetical prote
10	479	23.3	376	2	T24223	hypothetical prote
11	468	22.7	344	2	B89130	protein F52E1.8 [i
12	375	18.2	449	2	T15933	hypothetical prote
13	329	16.0	452	2	T20556	hypothetical prote
14	310	15.0	755	2	T19118	acid phosphatase h
15	287.5	14.0	380	2	T16883	hypothetical prote
16	274	13.3	413	2	T18945	hypothetical prote
17	271.5	13.2	416	2	T16058	hypothetical prote
18	250.5	12.2	408	2	T20893	hypothetical prote
19	247.5	12.0	471	2	T18944	hypothetical prote
20	220	10.7	462	2	T40420	probable acid phos
21	189.5	9.2	523	2	T15649	hypothetical prote
22	180.5	8.8	391	2	T27918	hypothetical prote
23	169	8.2	468	2	A86233	hypothetical prote
24	167.5	8.1	465	2	JE0369	histidine acid pho
25	158	7.7	969	2	T27997	hypothetical prote
26	158	7.7	1413	2	D88844	protein ZK792.1 [i
27	144.5	7.0	701	2	T20892	hypothetical prote
28	143.5	7.0	537	2	S54770	secreted acid phos
29	143.5	7.0	888	2	T46726	secreted acid phos
30	134.5	6.5	251	2	T46373	hypothetical prote
31	129.5	6.3	463	2	T39929	thiamin-repressibl
32	129	6.3	413	2	F90773	periplasmic glucos
33	129	6.3	413	2	B85636	periplasmic glucos
34	121	5.9	413	2	JV0087	glucose-1-phosphat
35	115.5	5.6	737	2	T30795	hypothetical prote
36	115.5	5.6	737	2	F42508	E2L protein - vacc
37	112.5	5.5	737	2	A72156	C2L protein - vari
38	110	5.3	599	1	RRXBPM	RNA-directed RNA p
39	109.5	5.3	737	2	E36841	E2L protein - vari
40	108.5	5.3	737	2	T28481	hypothetical prote
41	106	5.1	463	2	S14119	acid phosphatase (
42	105	5.1	413	2	AG0632	glucose-1-phosphat
43	105	5.1	417	2	S25627	glucose-1-phosphat
44	104.5	5.1	432	2	B36733	acid phosphatase (
45	104	5.0	489	1	RGBYP4	PET494 protein pre

ALIGNMENTS

RESULT 1
JH0152
acid phosphatase (EC 3.1.3.2) precursor, prostatic - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 31-Dec-1991 #sequence_revision 31-Dec-1991 #text_change 22-Jun-1999
C;Accession: JH0152; S41252
R;Roiko, K.; Jaenne, O.A.; Vihko, P.
Gene 89, 223-229, 1990
A;Title: Primary structure of rat secretory acid phosphatase and comparison to other
A;Reference number: JH0152; MUID:90323620
A;Accession: JH0152
A;Molecule type: mRNA
A;Residues: 1-381 <ROI>
A;Cross-references: GB:M32397; NID:g206028; PIDN:AAA41806.1; PID:g206029
R;Virkkunen, P.H.; Hedberg, P.; Palvimo, J.J.; Birr, E.; Porvari, K.; Taavitsainen, P
submitted to the EMBL Data Library, September 1993
A;Description: Structural organization of human and rat prostate-specific acid phosph
equence in the human gene promoter.
A;Reference number: S41251
A;Accession: S41252
A;Molecule type: DNA
A;Residues: 1-39 <VIR>
A;Cross-references: EMBL:X74969; NID:g439674; PIDN:CAA52914.1; PID:g439675
C;Superfamily: mammalian acid phosphatase
C;Keywords: glycoprotein; phosphohistidine; phosphoprotein; phosphoric monoester hydr
F;1-31/Domain: signal sequence #status predicted <SIG>
F;32-381/Product: acid phosphatase, prostatic #status predicted <MAT>
F;42/Active site: Arg #status predicted
F;43/Active site: His (phosphohistidine intermediate) #status predicted
F;93,219,332/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 88.3%; Score 1818; DB 2; Length 381;
Best Local Similarity 88.5%; Pred. No. 9.8e-140;
Matches 337; Conservative 24; Mismatches 20; Indels 0; Gaps 0;

```
Qy       1 MGAVPLPLSPTASLSLGFLLLLSLCLDPGQAKELKFVTLVFRHGDRGPIETFPTDPITES 60
           | |||| |  ||||:||||||||| ||||||||||||||||||||||||||||| ||| ||
Db       1 MRAVPLHLVGTASLTLGFLLLLSLRLDPGQAKELKFVTLVFRHGDRGPIETFPNDPIKES 60

Qy      61 SWPQGFGQLTQWGMEQHYELGSYIRKRYGRFLNDTYKHDQIYIRSTDVDRTLMSAMTNLA 120
           ||||||||||:||| ||||||||||:|||||||::|||||:||||||||||||||||||
Db      61 SWPQGFGQLTKWGMGQHYELGSYIRRRYGRFLNNSYKHDQVYIRSTDVDRTLMSAMTNLA 120

Qy     121 ALFPPEGISIWNPRLLWQPIPVHTVSLSEDRLLYLPFRDCPRFEELKSETLESEEFLKRL 180
           ||||||||||||||||||||||||||||||||||||||||||:|||||||:||||||||
Db     121 ALFPPEGISIWNPRLLWQPIPVHTVSLSEDRLLYLPFRDCPRFQELKSETLKSEEFLKRL 180

Qy     181 HPYKSFLDTLSSLSGFDDQDLFGIWSKVYDPLFCESVHNFTLPSWATEDAMIKLKELSEL 240
           |||||:||| |||||:||||| |||::||||:||||||||  :|||||| ||||||||
Db     181 QPYKSFIDTLPSLSGFEDQDLFEIWSRLYDPLYCESVHNFTFRTWATEDAMTKLKELSEL 240
```

```
Qy    241 SLLSLYGIHKQKEKSRLQGGVLVNEILKNMKLATQPQKYKKLVMYSAHDTTVSGLQMALD 300
          ||||||||||||||||||||||||||||||||||| :||:||||:||||||||||:
Db    241 SLLSLYGIHKQKEKSRLQGGVLVNEILKNMKLATQPQKARKLIMYSAYDTTVSGLQMALE 300

Qy    301 VYNGVLPPYASCHMMELYHDKGGHFVEMYYRNETQNEPYPLTLPGCTHSCPLEKFAELLD 360
          :|||:|||||||||:|||| | || ||||||||||||||||||||||||||||||||||||
Db    301 LYNGLLPPYASCHIMELYQDNGGTFVEMYYRNETQNEPYPLTLPGCTHSCPLEKFAELLD 360

Qy    361 PVIPQDWATECMATSSHQGTV 381
          |||||||||||| ||:|| ::
Db    361 PVIPQDWATECMGTSNHQASL 381


RESULT   2
JH0610
acid phosphatase (EC 3.1.3.2) ACPP precursor [validated] - human
N;Alternate names: acid phosphatase, prostatic; orthophosphoric monoester phosphohydrola
C;Species: Homo sapiens (man)
C;Date: 17-Aug-1992 #sequence_revision 01-Dec-1995 #text_change 08-Dec-2000
C;Accession: JH0610; JS0693; A38608; S01331; A32419; S11147; S38863; S41251; S17042; S42
R;Sharief, F.S.; Li, S.S.L.
Biochem. Biophys. Res. Commun. 184, 1468-1476, 1992
A;Title: Structure of human prostatic acid phosphatase gene.
A;Reference number: JH0610; MUID:92272747
A;Accession: JH0610
A;Molecule type: DNA
A;Residues: 1-386 <SHA>
A;Cross-references: GB:M97580; GB:M97581; GB:M97582; GB:M97583; GB:M97584; GB:M97585; GB
A;Accession: JS0693
A;Molecule type: mRNA
A;Residues: 1-386 <SH3>
A;Cross-references: GB:M97589; NID:g189611; PIDN:AAA60021.1; PID:g189613
R;Van Etten, R.L.; Davidson, R.; Stevis, P.E.; MacArthur, H.; Moore, D.L.
J. Biol. Chem. 266, 2313-2319, 1991
A;Title: Covalent structure, disulfide bonding, and identification of reactive surface a
A;Reference number: A38608; MUID:91115848
A;Accession: A38608
A;Molecule type: mRNA
A;Residues: 1-386 <VAN>
A;Cross-references: GB:M34840; NID:g189620; PIDN:AAA69694.1; PID:g189621
A;Note: part of this sequence, including the amino end of the mature protein, was confir
R;Vihko, P.; Virkkunen, P.; Henttu, P.; Roiko, K.; Solin, T.; Huhtala, M.L.
FEBS Lett. 236, 275-281, 1988
A;Title: Molecular cloning and sequence analysis of cDNA encoding human prostatic acid p
A;Reference number: S01331; MUID:88312981
A;Accession: S01331
A;Molecule type: mRNA
A;Residues: 1-14,'A',16,'ASC',20,'CF',23,'C',25-65,'WIWPTHPA',74-211,'A',213-386 <VIH>
A;Cross-references: EMBL:X52174; NID:g28321; PIDN:CAA36422.1; PID:g28322
A;Note: part of this sequence, including the amino end of the mature protein, was confir
R;Sharief, F.S.; Lee, H.; Leuderman, M.M.; Lundwall, A.; Deaven, L.L.; Lee, C.; Li, S.S.
Biochem. Biophys. Res. Commun. 160, 79-86, 1989
A;Title: Human prostatic acid phosphatase: cDNA cloning, gene mapping and protein sequen
A;Reference number: A32419; MUID:89228054
A;Accession: A32419
A;Molecule type: mRNA
A;Residues: 1-14,'A',16,'ASC',20,'CF',23,'C',25-94,'D',96-115,'R',117-214,'S',216-293,'T
A;Cross-references: GB:M24902; NID:g189618; PIDN:AAA60022.1; PID:g189619
A;Note: the authors translated the codons GAC for residue 95 as Glu, CGT for residue 116
R;Tailor, P.G.; Govindan, M.V.; Patel, P.C.
Nucleic Acids Res. 18, 4928, 1990
A;Title: Nucleotide sequence of human prostatic acid phosphatase determined from a full-
A;Reference number: S11147; MUID:90370491
A;Accession: S11147
A;Status: preliminary; translation not shown
A;Molecule type: mRNA
A;Residues: 1-14,'AFASC',20,'CF',23,'C',25-45,'H',47-65,'RIWPTHPA',74-138,'E',140-156,'R
A;Cross-references: EMBL:X53605; NID:g35683; PIDN:CAA37673.1; PID:g35684
R;Banas, B.; Blaschke, D.; Fittler, F.; Hoerz, W.
submitted to the EMBL Data Library, April 1993
A;Description: Characterization of the promoter of the human prostatic acid phosphatase
A;Reference number: S38863
A;Accession: S38863
A;Molecule type: DNA
A;Residues: 1-40 <BAN>
A;Cross-references: EMBL:X71391; NID:g416530; PIDN:CAA50514.1; PID:g416531
R;Virkkunen, P.H.; Hedberg, P.; Palvimo, J.J.; Birr, E.; Porvari, K.; Taavitsainen, P
submitted to the EMBL Data Library, September 1993
A;Description: Structural organization of human and rat prostate-specific acid phosph
equence in the human gene promoter.
A;Reference number: S41251
A;Accession: S41251
A;Molecule type: DNA
A;Residues: 1-40 <VIR>
A;Cross-references: EMBL:X74961; NID:g439657; PIDN:CAA52913.1; PID:g439658
R;Lee, H.; Chu, T.M.; Li, S.S.L.; Lee, C.
Biochem. J. 277, 759-765, 1991
A;Title: Homodimer and heterodimer subunits of human prostate acid phosphatase.
A;Reference number: S17042; MUID:91336999
A;Accession: S17042
A;Status: preliminary
A;Molecule type: protein
A;Residues: 33-49 <LEE>
R;Banas, B.; Blaschke, D.; Fittler, F.; Hoerz, W.
Biochim. Biophys. Acta 1217, 188-194, 1994
A;Title: Analysis of the promoter of the human prostatic acid phosphatase gene.
A;Reference number: S42730; MUID:94153995
A;Accession: S42730
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-40 <BA2>
A;Cross-references: GB:X71391; NID:g416530; PIDN:CAA50514.1; PID:g416531
R;Sharief, F.S.; Li, S.S.
Biochem. Mol. Biol. Int. 33, 561-565, 1994
A;Title: Nucleotide sequence of human prostatic acid phosphatase ACPP gene.
A;Reference number: I37175; MUID:95038536
A;Accession: I37175
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-386 <RES>
A;Cross-references: EMBL:U07097; NID:g515995; PIDN:AAB60640.1; PID:g515997
C;Comment: This protein is synthesized under androgen regulation by epithelial cells
C;Genetics:
A;Gene: GDB:ACPP
A;Cross-references: GDB:119644; OMIM:171790
A;Map position: 3q21.3-3q25.2
A;Introns: 40/3; 72/3; 101/3; 152/3; 185/3; 216/3; 260/3; 288/3; 323/2
C;Function:
A;Description: catalyzes the hydrolysis of a wide range of phosphate esters
C;Superfamily: mammalian acid phosphatase
C;Keywords: glycoprotein; phosphohistidine; phosphoprotein; phosphoric monoester hydr
F;1-32/Domain: signal sequence #status predicted <SIG>
F;33-386/Product: acid phosphatase ACPP #status experimental <MAT>
F;43,86/Active site: Arg #status predicted
F;44/Active site: His (phosphohistidine intermediate) #status predicted
F;94,220,333/Binding site: carbohydrate (Asn) (covalent) #status experimental
F;161-372,215-313,347-351/Disulfide bonds: #status experimental


  Query Match             80.7%;  Score 1661.5;  DB 1;  Length 386;
  Best Local Similarity   81.6%;  Pred. No. 5e-127;
  Matches  311;  Conservative   28;  Mismatches   41;  Indels    1;  Gaps    1;

Qy      1 MGAVPLPLSPTASLSLGFLLLLSLCLDPG-QAKELKFVTLVFRHGDRGPIETFPTDPITE 59
          | | || |:  |||||||| ||   ||    ||||||||||||||||| ||:||||||| |
Db      1 MRAAPLLLARAASLSLGFLFLLFFWLDRSVLAKELKFVTLVFRHGDRSPIDTFPTDPIKE 60

Qy     60 SSWPQGFGQLTQWGMEQHYELGSYIRKRYGRFLNDTYKHDQIYIRSTDVDRTLMSAMTNL 119
          |||||||||||| ||||||||| |||||| :|||::|||:|:|||||||||||||||||
Db     61 SSWPQGFGQLTQLGMEQHYELGEYIRKRYRKFLNESYKHEQVYIRSTDVDRTLMSAMTNL 120

Qy    120 AALFPPEGISIWNPRLLWQPIPVHTVSLSEDRLLYLPFRDCPRFEELKSETLESEEFLKR 179
          |||||||:|||||  |||||||||| ||||:|||||||:||||:|||:|||| ||
Db    121 AALFPPEGVSIWNPILLWQPIPVHTVPLSEDQLLYLPFRNCPRFQELESETLKSEEFQKR 180

Qy    180 LHPYKSFLDTLSSLSGFDDQDLFGIWSKVYDPLFCESVHNFTLPSWATEDAMIKLKELSE 239
```

```
          ||||| |: ||  |||   ||||||||||||||:||||||||||||||||| | ||:||||
Db    181 LHPYKDFIATLGKLSGLHGQDLFGIWSKVYDPLYCESVHNFTLPSWATEDTMTKLRELSE 240

Qy    240 LSLLSLYGIHKQKEKSRLQGGVLVNEILKNMKLATQPQKYKKLVMYSAHDTTVSGLQMAL 299
          |||||||||||||||||||||||||| :|| |||   ||||:|||||||||||||||||
Db    241 LSLLSLYGIHKQKEKSRLQGGVLVNEILNHMKRATQIPSYKKLIMYSAHDTTVSGLQMAL 300

Qy    300 DVYNGVLPPYASCHMMELYHDKGGHFVEMYYRNETQNEPYPLTLPGCTHSCPLEKFAELL 359
          |||||:||||||||: ||| :|| :|||||||||||:||||| |||||: |||||:||||:
Db    301 DVYNGLLPPYASCHLTELYFEKGEYFVEMYYRNETQHEPYPLMLPGCSPSCPLERFAELV 360

Qy    360 DPVIPQDWATECMATSSHQGT 380
           |||||||:|||| |:|||||
Db    361 GPVIPQDWSTECMTTNSHQGT 381


RESULT   3
S06167
acid phosphatase (EC 3.1.3.2) ACP2 precursor [validated] - human
N;Alternate names: acid phosphatase, lysosomal
C;Species: Homo sapiens (man)
C;Date: 04-Dec-1992 #sequence_revision 01-Dec-1995 #text_change 08-Dec-2000
C;Accession: S06167; S05525; S01155
R;von Figura, K.
submitted to the EMBL Data Library, June 1989
A;Reference number: S06167
A;Accession: S06167
A;Molecule type: DNA
A;Residues: 1-423 <VON>
A;Cross-references: EMBL:X15525; NID:g34239; PIDN:CAA33542.1; PID:g1199524
R;Geier, C.; von Figura, K.; Pohlmann, R.
Eur. J. Biochem. 183, 611-616, 1989
A;Title: Structure of the human lysosomal acid phosphatase gene.
A;Reference number: S05525; MUID:89377828
A;Accession: S05525
A;Molecule type: DNA
A;Residues: 1-29 <GEI>
R;Pohlmann, R.; Krentler, C.; Schmidt, B.; Schroeder, W.; Lorkowski, G.; Culley, J.; Mer
EMBO J. 7, 2343-2350, 1988
A;Title: Human lysosomal acid phosphatase: cloning, expression and chromosomal assignmen
A;Reference number: S01155; MUID:89052645
A;Accession: S01155
A;Molecule type: mRNA
A;Residues: 1-423 <POH>
A;Cross-references: EMBL:X12548; NID:g34262; PIDN:CAA31064.1; PID:g34263
A;Note: part of this sequence, including the amino end of the mature protein, was confir
C;Genetics:
A;Gene: GDB:ACP2
A;Cross-references: GDB:118963; OMIM:171650
A;Map position: 11p11.2-11p11.11
A;Introns: 38/3; 70/3; 99/3; 150/3; 183/3; 213/3; 258/1; 285/3; 321/2; 380/1
C;Function:
A;Description: catalyzes the hydrolysis of a wide range of phosphate esters
C;Superfamily: mammalian acid phosphatase
C;Keywords: glycoprotein; phosphohistidine; phosphoprotein; phosphoric monoester hydrola
F;1-30/Domain: signal sequence #status predicted <SIG>
F;31-423/Product: acid phosphatase ACP2 #status experimental <MAT>
F;41/Active site: Arg #status predicted
F;42/Active site: His (phosphohistidine intermediate) #status predicted
F;92,133,167,177,191,267,322,331/Binding site: carbohydrate (Asn) (covalent) #status pre
F;159-370,212-310,345-349/Disulfide bonds: #status predicted


  Query Match             49.6%;  Score 1022.5;  DB 1;  Length 423;
  Best Local Similarity   51.4%;  Pred. No. 4e-75;
  Matches  189;  Conservative   62;  Mismatches  112;  Indels    5;  Gaps    3;

Qy      14 LSLGFLLLLSLCLDPGQAKELKFVTLVFRHGDRGPIETFPTDPITESSWPQGFGQLTQWG 73
           | ||  |::   : | :|: |:||||::||||| |::|:| ||  |  |||||||||: |
Db      16 LLLGVNLVV---MPPTRARSLRFVTLLYRHGDRSPVKTYPKDPYQEEEWPQGFGQLTKEG 72

Qy      74 MEQHYELGSYIRKRYGRFLNDTYKHDQIYIRSTDVDRTLMSAMTNLAALFPPEGISIWNP 133
           | ||:|||  :|:||  ||| :|   ::|:|||| |||||||  ||| |||| |:  :||
Db      73 MLQHWELGQALRQRYHGFLNTSYHRQEVYVRSTDFDRTLMSAEANLAGLFPPNGMQRFNP 132

Qy     134 RLLWQPIPVHTVSLSEDRLLYLPFRDCPRFEELKSETLESEEFLKRLHPYKSFLDTLSSL 193
           : ||||||||| ::|||||  |   |||:|:|::|| :: |:          ||| :::
Db     133 NISWQPIPVHTVPITEDRLLKFPLGPCPRYEQLQNETRQTPEYQNESSRNAQFLDMVANE 192

Qy     194 SGFDDQDLFGIWSKVYDPLFCESVHNFTLPSWATEDAMIKLKELSELSLLSLYGIHKQKE 253
           :|  |  |  :|: |||  ||||  |   || ||:   | :|  | : |   |:||::| |
Db     193 TGLTDLTLETVWN-VYDTLFCEQTHGLRLPPWASPQTMQRLSRLKDFSFRFLFGIYQQAE 251

Qy     254 KSRLQGGVLVNEILKNMKLATQPQKYKKLVMYSAHDTTVSGLQMALDVYNGVLPPYASCH 313
           |:|||||||: :| ||: |     :  ||::|||||||:  ||||||||||   ||||||
Db     252 KARLQGGVLLAQIRKNLTLMATTSQLPKLLVYSAHDTTLVALQMALDVYNGEQAPYASCH 311

Qy     314 MMELYHDKGGHF-VEMYYRNETQNEPYPLTLPGCTHSCPLEKFAELLDPVIPQDWATECM 372
           : ||| :  |:| ||||:|||:   |:||:|||| | |||: |  | :||:|:||  ||
Db     312 IFELYQEDSGNFSVEMYFRNESDKAPWPLSLPGCPHRCPLQDFLRLTEPVVPKDWQQECQ 371

Qy     373 ATSSHQGT 380
           |    |
Db     372 LASGPADT 379


RESULT   4
S14742
acid phosphatase (EC 3.1.3.2) precursor - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 21-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 31-Mar-2000
C;Accession: S14742
R;Geier, C.; von Figura, K.; Pohlmann, R.
Biol. Chem. Hoppe-Seyler 372, 301-304, 1991
A;Title: Molecular cloning of the mouse lysosomal acid phosphatase.
A;Reference number: S14742; MUID:91282986
A;Accession: S14742
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-421 <GEI>
A;Cross-references: EMBL:X57199; NID:g52870; PIDN:CAA40485.1; PID:g52871
C;Superfamily: mammalian acid phosphatase
C;Keywords: lysosome; phosphoric monoester hydrolase


  Query Match             49.5%;  Score 1020.5;  DB 2;  Length 421;
  Best Local Similarity   51.4%;  Pred. No. 5.8e-75;
  Matches  187;  Conservative   66;  Mismatches  106;  Indels    5;  Gaps    3;

Qy      21 LLSLCL---DPGQAKELKFVTLVFRHGDRGPIETFPTDPITESSWPQGFGQLTQWGMEQH 77
           || :||    | ||: |:|||||::||||| |::|:| ||  |  |||||||||: || ||
Db      15 LLGMCLTVMPPIQARSLRFVTLLYRHGDRSPVKTYPKDPYQEEKWPQGFGQLTKEGMLQH 74

Qy      78 YELGSYIRKRYGRFLNDTYKHDQIYIRSTDVDRTLMSAMTNLAALFPPEGISIWNPRLLW 137
           :|||  :|:||  ||| :|   ::|:|||| |||||||  ||| ||||  :  ::| : |
Db      75 WELGQALRQRYHGFLNTSYHRQEVYVRSTDFDRTLMSAEANLAGLFPPNEVQHFSPNISW 134

Qy     138 QPIPVHTVSLSEDRLLYLPFRDCPRFEELKSETLESEEFLKRLHPYKSFLDTLSSLSGFD 197
           |||||||| ::|||||  |   |||:|:|::|| :: |:  |      ||: ::: :|
Db     135 QPIPVHTVPITEDRLLKFPLGPCPRYEQLQNETRQTPEYQNRSIQNAQFLNMVANETGLT 194

Qy     198 DQDLFGIWSKVYDPLFCESVHNFTLPSWATEDAMIKLKELSELSLLSLYGIHKQKEKSRL 257
           :  |  ||: ||| |||| |   || ||:   : :| :| : | | |:|||:| :|:||
Db     195 NVTLETIWN-VYDTLFCEQTHGLLLPPWASPQTVQRLSQLKDFSFLFLFGIHEQVQKARL 253

Qy     258 QGGVLVNEILKNMKLATQPQKYKKLVMYSAHDTTVSGLQMALDVYNGVLPPYASCHMMEL 317
           |||||: :|||||: |     :: ||::||||||:  |||||:||||   ||||||: ||
Db     254 QGGVLLAQILKNLTLMATTSQFPKLLVYSAHDTTLVALQMALNVYNGKQAPYASCHIFEL 313

Qy     318 YHDKGGHF-VEMYYRNETQNEPYPLTLPGCTHSCPLEKFAELLDPVIPQDWATECMATSS 376
           | :  |:| ||||:||:::  |:|| |||| | |||: |  | :||||:||  ||   :
Db     314 YQEDNGNFSVEMYFRNDSKKAPWPLILPGCPHRCPLQDFLRLTEPVIPKDWQKECQLAND 373

Qy     377 HQGT 380
           |
Db     374 TADT 377
```

```
RESULT   5
A33395
acid phosphatase (EC 3.1.3.2) precursor - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 29-Jan-1990 #sequence_revision 29-Jan-1990 #text_change 22-Jun-1999
C;Accession: A33395
R;Himeno, M.; Fujita, H.; Noguchi, Y.; Kono, A.; Kato, K.
Biochem. Biophys. Res. Commun. 162, 1044-1053, 1989
A;Title: Isolation and sequencing of a cDNA clone encoding acid phosphatase in rat liver
A;Reference number: A33395; MUID:89350910
A;Accession: A33395
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-423 <HIM>
A;Cross-references: GB:M27893; NID:g202933; PIDN:AAA40744.1; PID:g202934
C;Superfamily: mammalian acid phosphatase
C;Keywords: phosphohistidine; phosphoprotein; phosphoric monoester hydrolase
F;41/Active site: Arg #status predicted
F;42/Active site: His (phosphohistidine intermediate) #status predicted


  Query Match             49.5%;  Score 1019.5;  DB 2;  Length 423;
  Best Local Similarity   51.9%;  Pred. No. 7.1e-75;
  Matches  189;  Conservative   61;  Mismatches  109;  Indels    5;  Gaps    3;

Qy      21 LLSLCL---DPGQAKELKFVTLVFRHGDRGPIETFPTDPITESSWPQGFGQLTQWGMEQH 77
           || :||    | ||: |:||||::||||| |:: :| ||  |  |||||||||: || ||
Db      17 LLGMCLMVMPPIQARSLRFVTLLYRHGDRSPVKAYPKDPYQEEKWPQGFGQLTKEGMLQH 76

Qy      78 YELGSYIRKRYGRFLNDTYKHDQIYIRSTDVDRTLMSAMTNLAALFPPEGISIWNPRLLW 137
           :|||  :|:||  ||| :|   ::|:|||| |||||||  ||| |||  :  :|| : |
Db      77 WELGQALRQRYHGFLNASYHRQEVYVRSTDFDRTLMSAEANLAGLFPPTEVQHFNPNISW 136

Qy     138 QPIPVHTVSLSEDRLLYLPFRDCPRFEELKSETLESEEFLKRLHPYKSFLDTLSSLSGFD 197
           |||||||| ::|||||  |   |||:|:|::|| :: |:         ||| ::: :|
Db     137 QPIPVHTVPITEDRLLKFPLGPCPRYEQLQNETRQTPEYQNMSIQNAQFLDMVANETGLM 196

Qy     198 DQDLFGIWSKVYDPLFCESVHNFTLPSWATEDAMIKLKELSELSLLSLYGIHKQKEKSRL 257
           :  |  ||: ||| |||  |   || ||:   :  | :| : | | |:||| | :|:||
Db     197 NLTLETIWN-VYDTLFCEQTHGLLLPPWASPQTVQALSQLKDFSFLFLFGIHDQVQKARL 255

Qy     258 QGGVLVNEILKNMKLATQPQKYKKLVMYSAHDTTVSGLQMALDVYNGVLPPYASCHMMEL 317
           |||||: :|||: |     :: ||::|||||||:  |||||:||||   ||||||: ||
Db     256 QGGVLLAQILKNLTLMATTSQFPKLLVYSAHDTTLVALQMALNVYNGKQAPYASCHIFEL 315

Qy     318 YHDKGGHF-VEMYYRNETQNEPYPLTLPGCTHSCPLEKFAELLDPVIPQDWATECMATSS 376
           | :  |:| ||||:||:::  |:|||||| | |||: |  | :||||:||  ||   |
Db     316 YQEDNGNFSVEMYFRNDSKKAPWPLTLPGCPHRCPLQDFLRLTEPVIPKDWQKECQLASD 375

Qy     377 HQGT 380
             |
Db     376 TADT 379


RESULT   6
S64682
acid phosphatase (EC 3.1.3.2) Acph-1 precursor, lysosomal - fruit fly (Drosophila melanc
C;Species: Drosophila melanogaster
C;Date: 06-Dec-1996 #sequence_revision 06-Dec-1996 #text_change 17-Mar-2000
C;Accession: S64682; S64681
R;Chung, H.J.; Shaffer, C.; MacIntyre, R.
Mol. Gen. Genet. 250, 635-646, 1996
A;Title: Molecular characterization of the lysosomal acid phosphatase from Drosophila me
A;Reference number: S64681; MUID:96194627
A;Accession: S64682
A;Status: not compared with conceptual translation
A;Molecule type: DNA
A;Residues: 1-438 <CHU>
A;Accession: S64681
A;Status: not compared with conceptual translation
A;Molecule type: mRNA
A;Residues: 1-31,'Y',33-438 <CHW>
C;Genetics:
A;Gene: FlyBase:Acph-1
A;Cross-references: FlyBase:FBgn0000032
C;Superfamily: mammalian acid phosphatase
C;Keywords: phosphoric monoester hydrolase
F;1-33/Domain: signal sequence #status predicted <SIG>
F;34-438/Product: acid phosphatase Acph-1 #status predicted <MAT>


  Query Match             30.0%;  Score 617;  DB 2;  Length 438;
  Best Local Similarity   38.1%;  Pred. No. 3.4e-42;
  Matches  141;  Conservative   59;  Mismatches  114;  Indels   56;  Gaps   14;

Qy      28 PGQAKELKFVTLVFRHGDRGPIETFPTDPITESS-WPQGFGQLTQWGMEQHYELGSYIRK 86
           |||   |||| :::|||| |:: :|||  :   || |:| ||  | ::||:|| ::|
Db      45 PGQ---LKFVHVIYRHGDRTPVDPYPTDPWGDRKFWPTGWGDLTNLGKQEHYDLGKWLRN 101

Qy      87 RYGRFLNDTYKHDQIYIRSTDVDRTLMSAMTNLAALFPPEGISIWNPRLLWQPIPVHTVS 146
           ||   |  | :: ||::|||||||||||| :||| |: |:|  |||  : |||||:||
Db     102 RYSNLLPPIYSNENIYVQSTDVDRTLMSAQSNLAGLYEPQGEDIWNTDINWQPIPIHTSP 161

Qy     147 LSEDRLLYLPFRDCPRFE-ELKSETLESEEFLKRL----------------HPYKSFLDT 189
             || :|      || :: || |  |||    | |                  | |:|:|
Db     162 EREDPILAAK-APCPAYDYELAS--LESSPEFKALTEKHRNLFAYLSEKGGRPVKTFIDA 218

Qy     190 LSSLSGFDDQDLFGIWSKVYDPLFCESVHNFTLPSWATEDAMIKLKELSELSLLS--LYG 247
                              : : || |:::| ||| |       |:    ||: :|   :
Db     219 ---------------QYLNNTLFIENLYNMTLPKWTK-----KVYGREELTYVSNFAFA 257

Qy     248 IHKQKEK-SRLQGGVLVNEILKNMKL----ATQPQKYKKLVMYSAHDTTVSGLQMALDVY 302
           |     | :||: | |: :| :  |     : :|   : : :|||||||: :  || ::
Db     258 ISSYTRKLARLKAGPLLKDIFQRFKEKSSGSLKPD--RSMWVYSAHDTTVASVLNALKLF 315

Qy     303 NGVLPPYASCHMMELYHDK-GGHFVEMYYRNETQNEPYPLTLPGCTHSCPLEKFAELLDP 361
              ||| :| ||||  |:     | ::|:| |  || || :|||  |||| |   : :
Db     316 ELHSPPYTACIMMELRVDETNTPLVSIFYKNTTA-EPLPLDIPGCGPSCPLTKLMNIYED 374

Qy     362 VIPQDWATEC 371
           |:| ||  ||
Db     375 VLPVDWEREC 384


RESULT   7
D88504
protein B0361.7 [imported] - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 24-May-2001
C;Accession: D88504
R;anonymous, The C. elegans Sequencing Consortium.
Science 282, 2012-2018, 1998
A;Title: Genome sequence of the nematode C. elegans: a platform for investigating bio
A;Reference number: A75000; MUID:99069613; PMID:9851916
A;Note: see websites genome.wustl.edu/gsc/C_elegans/ and www_sanger.ac.uk/Projects/C_
A;Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999;
A;Accession: D88504
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-411 <STO>
A;Cross-references: GB:chr_III; PIDN:AAA50626.1; PID:g458955; GSPDB:GN00021; CESP:B03
C;Genetics:
A;Gene: B0361.7
A;Map position: 3
C;Superfamily: mammalian acid phosphatase


  Query Match             26.3%;  Score 542.5;  DB 2;  Length 411;
  Best Local Similarity   34.7%;  Pred. No. 3.4e-36;
  Matches  135;  Conservative   61;  Mismatches  138;  Indels   55;  Gaps   14;

Qy      19 LLLLSLCLD-----------PGQAK--ELKFVTLVFRHGDRGPIETFPTDPITESSWPQ 64
           |||| | :             | ||    |::|  |:|||| | |    | ||:  ||:
Db       2 LLLLVLLIGASGINAVVYKEVPIQANTDTLEYVHTVWRHGDRTPAELLFPDDITK--WPE 59
```

```
Qy      65 GFGQLTQWGMEQHYELGSYIRKRYGRFLNDTYKHDQIYIRSTDVDRTLMSAMTNLAALFP 124
Db      60 GLGELTEQGAAQQYRLGQWLKRRYGSWLGEKFNRNAIYIRSSDYNRTLMSAQANMAGLFP 119

Qy     125 PE-----GISIWNPRLLWQPIPVHTVSLSEDRLLYLPFRDCPRFEELKSETLESE----- 174
Db     120 PKYPIAGG-------LMWQPIPVHTISKPTDKELY---------EEASCPTAEIEMNAQW 163

Qy     175 EFLKRLHPYKSFLDTLSSLS---GFDDQDLFGIWSKVYDPLFCESVHNFTLPSWATEDAM 231
Db     164 KSTKANGIRKKFARELSFFSQKLNLPNMELKATW-RIFDNLFCEKQNNITWPSWMNSSIF 222

Qy     232 IKLKEL-SELSLLSLYGIHKQKEKSRLQGGVLVNEILKNM--KLATQPQKYKKLVMYSAH 288
Db     223 ERVDQLYNEVSQLEFH----TDTLRRLRGGTLLEEIFHRFSDKASGSLGKEAKFYAYSAH 278

Qy     289 DTTVSGLQMALDVYNGVLPPYASCHMMELYH-DKGGHFVEMYYRNETQ-NEPYPLTLPGC 346
Db     279 DSTIAALLATLGVFYDIYPKYATCLLIEMHKLANETRLIRVFHKNETDIDRLIEYSIPGC 338

Qy     347 THSCPLEKFAELLDPVIPQDWATECMATS 375
Db     339 DDPCTLQKLGDDLKKYFPEDWEAECVIIS 367
```

RESULT 8
T32457
hypothetical protein T13B5.3 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 17-Mar-2000
C;Accession: T32457
R;Maggi, L.
submitted to the EMBL Data Library, September 1997
A;Description: The sequence of C. elegans cosmid T13B5.
A;Reference number: Z21172
A;Accession: T32457
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-440 <MAG>
A;Cross-references: EMBL:AF026211; PIDN:AAB71294.1; GSPDB:GN00020; CESP:T13B5.3
A;Experimental source: strain Bristol N2; clone T13B5
C;Genetics:
A;Gene: CESP:T13B5.3
A;Map position: 2
A;Introns: 25/3; 90/3; 125/3; 162/2; 242/3; 289/1; 328/3; 373/3; 404/3
C;Superfamily: mammalian acid phosphatase

Query Match 23.6%; Score 487; DB 2; Length 440;
Best Local Similarity 33.8%; Pred. No. 1.2e-31;
Matches 130; Conservative 61; Mismatches 144; Indels 50; Gaps 14;

```
Qy      33 ELKFVTLVFRHGDRGPIETFPTDPITESSWPQGFGQLTQWGMEQHYELGSYIRKRY---- 88
Db      52 ELLLSQVVWRHGDRAPTGTYPTDPHKEEAWPNGWGELTQLGMRQQYALGRLLYKKYVNST 111

Qy      89 --GRFLNDTYKHDQIYIRSTDVDRTLMSAMTNLAALFP--------PEGISIWNPRLLWQ 138
Db     112 GPSKLLSSSYNSKEVYIRSTDVNRTLVSALANLAGMFENGNRGADYPDS-KRWPTN--WT 168

Qy     139 PIPVHTVSLSEDRLLYLPFRDCPRFEELKSETLESEEFLKRLHPYKSFLDTLSSLSGFDD 198
Db     169 PIPIHTLAEKDDPVGNV-FAPCARAEELTRQIYAGSGFQKFVAENQEFLDFVSEKSGKK- 226

Qy     199 QDLFGIWSKVY---DPLFCESVHNFTLPSWATEDAMIKLKELSELSLLSLYGIHKQ--KE 253
Db     227 ----VIMPEIYMVNDVHFIEKLYNMSQPEWITDDVEIKLRNLSQVSTRFLFGIGDPYIPE 282

Qy     254 KSRLQGGVLVNEIL--KNMKLATQPQKYK----------KLVMYSAHDTTVSGLQMALD- 300
Db     283 LIRLRGGPLLRAMMDKMNQKLSCLAKNNEGEDCSWIGKLKYHAYSAHDTTVYAFLTTFGD 342

Qy     301 ---VYNGVLPPYASCHMMELYHDK-GGHFVEMYYRNETQNEPYPLT--LPGCTHS---CP 351
Db     343 EERVIEGGMPHYTASVAVELWNLKNGGPSVRVLFHSAFHHNYHVITHLAKGCPHNSEFCP 402

Qy     352 LEKFAELLDPVIPQDWATECMATSS 376
Db     403 LKTFEQRSLKFLPVNLQKECAPKKS 427
```

RESULT 9
T25060
hypothetical protein T21B6.2 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 18-Feb-2000
C;Accession: T25060
R;Cottage, A.
submitted to the EMBL Data Library, November 1995
A;Reference number: Z19975
A;Accession: T25060
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-366 <WIL>
A;Cross-references: EMBL:Z68011; PIDN:CAA92013.1; GSPDB:GN00028; CESP:T21B6.2
A;Experimental source: clone T21B6
C;Genetics:
A;Gene: CESP:T21B6.2
A;Map position: X
A;Introns: 57/3; 86/3; 138/3; 168/2; 194/3; 239/2; 268/3; 305/3
C;Superfamily: mammalian acid phosphatase

Query Match 23.3%; Score 480; DB 2; Length 366;
Best Local Similarity 32.0%; Pred. No. 3.4e-31;
Matches 116; Conservative 77; Mismatches 129; Indels 40; Gaps 12;

```
Qy      18 FLLLLSLCLDPGQAKELKFVTLVFRHGDRGPIE-TFPTDPITESSWPQGFGQLTQWGMEQ 76
Db       5 YLFCCSLLLP--CAAEIEFLLAVWRHGDRAPENLPYPSDPHNETFWPRGWNQLTNVGIDQ 62

Qy      77 HYELGSYIRKRYGRFLNDTYKHDQIYIRSTDVDRTLMSAMTNLAALFPPEGISIWNPR-- 134
Db      63 ATKLGKFLRRRYQGSVLPVFDRKKISIRSSDADRAIETAQSVATALFPPDGLQVWNEEKF 122

Qy     135 LLWQPIPVHTVSLSEDRLLYLPFRDCPRFEELKSE---TLESEEFLKRLHPYKSFLDTLS 191
Db     123 RFWQPIPIRT-NGKPDPMLRPSKIQCPAYQRIVAEERKKIESEINVK----YKRELEIIS 177

Qy     192 SLSGFDDQDLFGIWSKVYDPLFCESVHNFTLPSWATE--------DAMIKLKELSELSLL 243
Db     178 NHT--SHQTKYGNIKDVYN-----HYNGLPFPNWIDEKVNGKSLLDTIAEIRRIARLQLF 230

Qy     244 SLYGIHKQKEKSRLQGGVLVNEILKNMKLATQPQKYKKLVMYSAHDTTVSGLQMALDVYN 303
Db     231 N------SRAKAKFMAGYLINSWTESLVLASQHISPKKALLYSSHDGTLSALMYGLGISN 284

Qy     304 GVLPPYASCHMMELYHDKGGHFVEMYYRNETQNEP---YPLTLPGCTHSCPLEKFAELLD 360
Db     285 HQLIPYTACIMIELH---TGNNVKIYFRNTTTENPDDVHEMFVPGCSTDCHLSKFIKSVN 341

Qy     361 PV 362
Db     342 GV 343
```

RESULT 10
T24223
hypothetical protein R13H4.3 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 21-Jan-2000
C;Accession: T24223
R;Kershaw, J.
submitted to the EMBL Data Library, November 1996
A;Reference number: Z19858

```
A;Accession: T24223
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-376 <WIL>
A;Cross-references: EMBL:Z81579; PIDN:CAB04655.1; GSPDB:GN00023; CESP:R13H4.3
A;Experimental source: clone R13H4
C;Genetics:
A;Gene: CESP:R13H4.3
A;Map position: 5
A;Introns: 59/3; 90/3; 144/3; 284/3; 324/3
C;Superfamily: mammalian acid phosphatase


  Query Match             23.3%;  Score 479;  DB 2;  Length 376;
  Best Local Similarity   31.1%;  Pred. No. 4.3e-31;
  Matches  108;  Conservative   72;  Mismatches  137;  Indels   30;  Gaps    9;

Qy      30 QAKELKFVTLVFRHGDRGPIE-TFPTDPITESSWPQGFGQLTQWGMEQHYELGSYIRKRY 88
           ::::|:||  ::|||||  |:   :| |  || :| :|:||||  ||:| :||| : | :|
Db      17 ESRKLEFVQALWRHGDRAPLHLPYPNDQYTEKAWSRGWGQLTSIGMQQLHELGEFFRHQY 76

Qy      89 --GRFLNDTYKHDQIYIRSTDVDRTLMSAMTNLAALFPPEGISIWNPRLLWQPIPVHTVS 146
             |:   :  ::|:||:| || |:||   |  |:|  |   |:  : |||:|||  :
Db      77 VDSSFIPSNFSVKEVYLRSSDSDRALVSAQAFLYGLYPASGGYQWSSDIDWQPLPVHAST 136

Qy     147 LSEDRLLYL---PFRDCPRFEELKSE-TLESEEFLKRLHPYKSFLDTLSSLSGF------ 196
           |  |: :       | | | | ::   ||  :      |  |   ||  :||
Db     137 PGEPDLVTVCKPTAIKCARHEVLVAQGDQESNAYYS--VKYADFFSELSQTTGFKHCSYM 194

Qy     197 DDQDLFGIWSKVYDPLFCESVHNFTL--PSWATE-----DAMIKLKELSELSLLSLYGIH 249
           |   || |          | :|| |   | | |:     |    :  ::|:  : :  :
Db     195 DINGLFDIQR--------ELIHNMTAKQPYWVTQTWPQYDNRTSMDIITEMRTVRMMNLF 246

Qy     250 KQKEKSRLQGGVLVNEILKNMKLATQPQKYKKLVMYSAHDTTVSGLQMALDVYNGVLPPY 309
             :|| :|:|| ::   ::|    :: :  :::::||:||  :  |  |    | :: ||
Db     247 NSEEKGKLEGGSVLYNWIQNAIAVSESRNDQRMLLYSSHDGVLLALLNAFRASNEMMVPY 306

Qy     310 ASCHMMELYHDKGGHFVEMYYRNETQNEPYPLTLPGCTHSCPLEKFA 356
           |:  :| :| | |  : |:|||||| ::|| : |  |   | : : |
Db     307 AAALIMHVYSDNGKFYSELYYRNETTSDPYRIPLSRCPEPCEVSQLA 353


RESULT  11
B89130
protein F52E1.8 [imported] - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 10-May-2001
C;Accession: B89130
R;anonymous, The C. elegans Sequencing Consortium.
Science 282, 2012-2018, 1998
A;Title: Genome sequence of the nematode C. elegans: a platform for investigating biolog
A;Reference number: A75000; MUID:99069613; PMID:9851916
A;Note: see websites genome.wustl.edu/gsc/C_elegans/ and www_sanger.ac.uk/Projects/C_ele
A;Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and
A;Accession: B89130
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-344 <STO>
A;Cross-references: GB:chr_V; PIDN:AAB37044.1; PID:g1086806; GSPDB:GN00023; CESP:F52E1.8
C;Genetics:
A;Gene: F52E1.8
A;Map position: 5


  Query Match             22.7%;  Score 468;  DB 2;  Length 344;
  Best Local Similarity   33.2%;  Pred. No. 3e-30;
  Matches  109;  Conservative   69;  Mismatches  138;  Indels   12;  Gaps    8;

Qy      40 VFRHGDRGPIE-TFPTDPITESSWPQGFGQLTQWGMEQHYELGSYIRKRYGRFLNDTYKH 98
           ::|||||  | :  :| |   |: ||:|: |||  |: |  ||| ::|:|||  :   :
Db       4 IWRHGDRAPGDLPYPKDKYNETFWPRGWDQLTNKGIWQAVELGIWLRQRYGATVLPIFDK 63

Qy      99 DQIYIRSTDVDRTLMSAMTNLAALFPPEGISIWNPRLL--WQPIPVHTVSLSEDRLLYLP 156
```

```
           |:::| |:| :| : :|    | ||||    :|    |  ||| |: |   : | ||
Db      64 DKVFILSSDSERAIETAQGVSAGLFPPVDDRVWESSYLRYWQPTPIQTAYGTIDALLRPT 123

Qy     157 FRDCPRFEELKSETLESEEFLKRLHPYKSFLDTLSSLSGFDDQDLFGIWSKVYDPLFCES 216
             || : :| :|  |:     :  : |    : | :::| :  | : | : :|| :  |
Db     124 KVKCPNY-DLANEQEEAPIATQINNEYGQMFNWLQNITGMESIDFWNI-NDLYD-IQREL 180

Qy     217 VHNFTLPSWATE--DAMIKLKELSELSLLSLYGIHKQKEKSRLQGGVLVNEILKNMKLAT 274
           ||   | |  :  :    :  : ||  ::         |:: :||:|||: |:||:
Db     181 DHNMPQPQWLNQVFNGTTIMDHIRELKRITRNQEFNSPTKAKFRGGMLVNQFLQNMEDLK 240

Qy     275 QPQKYKKLVMYSAHDTTVSGLQMALDVYNGVLPPYASCHMMELYHDKGGHFVEMYYRNET 334
             :  |  :|||:|| |:| |  ||:| |  | || :  : ||| |   : |:::|:| |
Db     241 ANKTTKNAMMYSSHDGTLSALLYALNVSNDQLVPYTATVLFELYDD---NTVQLFYKN-T 296

Qy     335 QNEPYPLTLPGCTHSCPLEKFAELLDPV 362
           :  ||:|:|||   ||  :| :||: |
Db     297 TSTAYPMTIPGCQQICPYSQFLQLLENV 324


RESULT  12
T15933
hypothetical protein EGAP2.3 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 18-Feb-2000
C;Accession: T15933
R;Bentley, D.
submitted to the EMBL Data Library, December 1995
A;Description: The sequence of C. elegans cosmid EGAP2.
A;Reference number: Z18430
A;Accession: T15933
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-449 <BEN>
A;Cross-references: EMBL:U42842; NID:g1125787; PID:g1125790; PIDN:AAA83593.1; CESP:EG
C;Genetics:
A;Gene: CESP:EGAP2.3
A;Introns: 32/3; 67/3; 99/3; 159/3; 191/3; 222/3; 262/1; 319/3; 415/3; 437/1
C;Superfamily: mammalian acid phosphatase


  Query Match             18.2%;  Score 375;  DB 2;  Length 449;
  Best Local Similarity   29.5%;  Pred. No. 1.5e-22;
  Matches  127;  Conservative   66;  Mismatches  138;  Indels  100;  Gaps   21;

Qy      32 KELKFVTLVFRHGDRGPIETFPTDPITESSWP---QGFGQLTQWGMEQHYELGSYIRKRY 88
           |:| ||  ::|||||  | :|| |||  | :|     |:|||:  ||:||  ||  :| ||
Db      25 KDLVFVQTLWRHGDRSPTKTFKTDPFQEDAWQFGGGGWGQLSPAGMKQHLNLGKMLRNRY 84

Qy      89 ---GRFLNDTYKHDQIYIRSTDVDRTLMSAMTNLAALFP-----------------PEG 127
             || : |   |||:|||||:||::|||:||   :                     | |
Db      85 VTNYNFLPNKYNAKQIYVRSTDVNRTIISAMSNLLGQYGQNDNSSTPGLDYPDVDGWPAG 144

Qy     128 ISIWNPRLLWQPIPVHTVSLSEDRLLYL----PFRDCPRFEELKSETLESEEFLKRLHPY 183
                 : || ||||    | |  :    ||:|            :  |  |
Db     145 ---------YVPIAVHTVDDDTDHLGNMESTCPFKD-----------QVWELAKTSDEV 183

Qy     184 KSFLDT------LSSLSGFDDQ--DLFGIWSKVYDPLFCESV-HNFTL---PSWATEDAM 231
           |||:::       | :|: :  |  |:  :|  : : |: | : :| ||    :| |:
Db     184 KSFVNSADVQAVLGNLTNYCGQPVDIDNLWI-ITNALYIEQIYYNATLRTKNNWFTDAFY 242

Qy     232 IKLKELSELSLLSLYGIHKQKEK----------SRLQGGVLVNEILK--NMKLATQPQ- 277
           |   :::   |   || |                :::|| ::|:::   |:||  | |
Db     243 AKADAINDQVQLFQNGIFKTVPNIVNGHDVGVLTRKVRGGPILNDMVMHINLKLMCQGQT 302

Qy     278 -------KYKKLVMYSAHDTTVSGLQMAL--DVY----NGVLPPYASCHMMELY---HDK 321
                  |  :|||||||:     ||  : |    :|  | |::  ::|||    ||
Db     303 TPNCTWINNLKNYIYSAHDTTIYAFFSALLIEEYAVKPSGGYPLYSAAVLLELYIDSVDK 362

Qy     322 GGHFVEMYYRNETQN-EPYPLTLPGCTHS---CPLE---KFAELLDPVIPQD-WATECMA 373
             :|  :|:  :    :   : : ||  :   | |:    ||  : |  | | |    :
Db     363 KPYFKMVYHEQDGSGFKDVTMGIQGCPQNSSYCDLDILRNFANTIKPDQPIDQWCLTDLN 422
```

```
Qy    374 TSSHQGTVGAL 384
          ||   ||  |
Db    423 KSSSFATVSML 433


RESULT  13
T20556
hypothetical protein F07H5.9 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 04-Mar-2000
C;Accession: T20556
R;Steward, C.
submitted to the EMBL Data Library, December 1995
A;Reference number: Z19292
A;Accession: T20556
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-452 <WIL>
A;Cross-references: EMBL:Z68314; PIDN:CAA92657.1; GSPDB:GN00020; CESP:F07H5.9
A;Experimental source: clone F07H5
C;Genetics:
A;Gene: CESP:F07H5.9
A;Map position: 2
A;Introns: 75/3; 171/3; 273/1; 331/3; 378/3
C;Superfamily: mammalian acid phosphatase


  Query Match             16.0%;  Score 329;  DB 2;  Length 452;
  Best Local Similarity   29.3%;  Pred. No. 8.3e-19;
  Matches  119;  Conservative   53;  Mismatches  150;  Indels   84;  Gaps   21;

Qy     40 VFRHGDRGPIETFPTDPITESSWP---QGFGQLTQWGMEQHYELGSYIRKRY-------G 89
          |:|||||  |  ||  ||  | ||     |:|||:  || ||  ||  :| ||
Db     41 VWRHGDRSPTTTFNVDPFQEDSWTFGGGGWGQLSPLGMNQHLTLGKKLRNRYVNTGNSTH 100

Qy     90 RFLNDTYKHDQIYIRSTDVDRTLMSAMTNLAALFPPEGISI-----------WNPRLLWQ 138
           ||   |    :|||||  ::|||:|| :|:  ::  :|               | ||  :
Db    101 NFLPAVYDQKTMYIRSTGINRTLISATSNMLGMYGQDGYGSSAGTDFPDAQGW-PR-GFV 158

Qy    139 PIPVHTVSLSEDRLLYLPFRDCPRFEELKSETLESEEF--LKRLHPYKSFLDTLSSLSGF 196
          |||||||    | :  :    ||| | | :   :|:|    :   |  | :| |:||
Db    159 PIPVHTVDYDSDHIGNMDC-ICPRREWLWNIAQQSDEMKNWRSSVPVSSVIDKLTSLVN- 216

Qy    197 DDQDLFGIWSKVYDPLFCESVH-NFTL---PSWATEDAMIKLKELSELSLLSLYGIHKQK 252
              |  |: | | || | :: | :|    :|  :|   ::  :::   :  ||:
Db    217 QTWSLEDFWT-VPDALFIEQIYFNESLRKENTWFNDDFYSQIVAVNDQIYMYQYGVFNST 275

Qy    253 ----------EKSRLQGGVLVNEILKNMKLATQPQKYK-----------KLVMYSAHDTT 291
                   |  ::: | |:|::  :   :     |           |  :|||| |
Db    276 INMNNMNIGLELLKVRSGPLMNDMMDRINKKSDCTYGKNTQGCDWINGLKYFVYSAHDET 335

Qy    292 VSGLQMALDVYNGVLPPYASCHMMELYHDKGGHFVEMYYRNETQNEPY------------ 339
          |  : :|| :    : |    |   ||       || |:|| | |  |
Db    336 VYAVLVALGIERFAIKP----HGYPLY--SAAVSVE-YWRNTTDNADYFKLVYHKQSGDD 388

Qy    340 ----PLTLPGCTHS-CP---LEKFAELLDPVIPQD-WATECMATSS 376
              : ||    |    | : |  | |  | | |    |: | |
Db    389 FTVMTKEITGCGGDYCSMNVLNQIANNLKPDQPIDQW---CLVTGS 431


RESULT  14
T19118
acid phosphatase homolog F26C11.1 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 29-Oct-1999
C;Accession: T19118; T21390
R;Lloyd, C.
submitted to the EMBL Data Library, October 1995
A;Reference number: Z19074
A;Accession: T19118
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-755 <WIL>
A;Cross-references: EMBL:Z54342; PIDN:CAA91156.1; GSPDB:GN00020; CESP:F26C11.1
A;Experimental source: clone C08H9
R;Matthews, P.
submitted to the EMBL Data Library, December 1994
A;Reference number: Z19416
A;Accession: T21390
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-755 <WI2>
A;Cross-references: EMBL:Z47072; PIDN:CAA87370.1; GSPDB:GN00020; CESP:F26C11.1
A;Experimental source: clone F26C11
C;Genetics:
A;Gene: CESP:F26C11.1
A;Map position: 2
A;Introns: 151/1; 199/3; 277/1; 316/2; 352/3; 417/3; 449/1; 475/3; 577/1; 635/3; 728/


  Query Match             15.0%;  Score 310;  DB 2;  Length 755;
  Best Local Similarity   26.7%;  Pred. No. 6e-17;
  Matches  112;  Conservative   65;  Mismatches  148;  Indels   94;  Gaps   16;

Qy     33 ELKFVTLVFRHGDRGPIETFPTDPITESSWP---QGFGQLTQWGMEQHYELGSYIRKRY- 88
          :|:||  ::|||||  :|     ||:| :|    | |:||  || :   ||:  |:||
Db    346 KLEFVQTIWRHGDRSALEGL--FPISEKNWTFGGGGLGELTPMGMSEMNNLGTIFRRRYV 403

Qy     89 --GRFLNDTYKHDQIYIRSTDVDRTLMSAMTNLAALFPPEGISIWNPR-------LLWQP 139
            :||:  |   :|||||:::||::|||: |  :|||     || :       : ||
Db    404 EDQQFLSHRYAAKEIYIRSTNLNRTIISAMSLLYGMFPP---GAWNIQGVDYPNDVDWQQ 460

Qy    140 ----IPVHTVSLSEDRLLYLPFRDCPRFEELKSE------------------------- 169
              ||||   :  |:       :| ||:||: :
Db    461 GFTFIPVHVDGI--DQCAVAQLCNCRRFQELQEKWAELDEVKNATVAMIALNRRVAAFYN 518

Qy    170 -TLESEEFLKRLHPYKS----FLDTLSSLSGFDDQDLFGIWSKVYDPLFCESVHNFTLPS 224
           | : |:| :    :|     | ||:     : ::||:    : | |    :  ||  |
Db    519 VTDQPEKFNRYTDAWKCQRNWFNDTMYQQLPWYNEDLYNEAQRTYAPFKRFTEGNFGNPK 578

Qy    225 WATEDAMIKLKELSELSLLSLYGIHKQKEKSRLQGGVLVNEILKNMK----------LA 273
           :  |                ||   :| | |||| |:||| :  :          :
Db    579 PSIVD----------------GIDIPQEVSTLQGGPLLNEIFERGREKIRCVADAENCS 621

Qy    274 TQPQKYKKLVMYSAHDTTVSGLQMALDVYNGV-----LPPYASCHMMELYHDKGGH-FVE 327
              |  |   ||:||  |  | : | : : |      |  :|   :| | : |    |:
Db    622 IDYLKPLKFYAYSSHDQLVYALLVTLGITDVVKTVDGWPDTSSSLTIEYYSNPGNQSSVK 681

Qy    328 MYYRNETQNEPYPLT--LPGCTHS--CPLEKFAELLDPVIP-QDWATECMATSSHQGTV 381
            ||: : :    :|  :| |  :  | :  |  :      |  |: | |  : |   :|
Db    682 FLYRDNSNDNFSDVTSQIPVCNGAQYCAMSDFQNIAYQFKPLPDYMTLCETSLSSISSV 740


RESULT  15
T16883
hypothetical protein T16D1.2 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 18-Feb-2000
C;Accession: T16883
R;Bentley, D.
submitted to the EMBL Data Library, October 1995
A;Description: The sequence of C. elegans cosmid T16D1.
A;Reference number: Z18597
A;Accession: T16883
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-380 <BEN>
A;Cross-references: EMBL:U39471; NID:g1041874; PID:g1041876; PIDN:AAA80132.1; CESP:T1
A;Experimental source: strain Bristol N2
C;Genetics:
A;Gene: CESP:T16D1.2
A;Introns: 34/3; 69/3; 137/3; 174/3; 278/1; 333/3
C;Superfamily: mammalian acid phosphatase


  Query Match             14.0%;  Score 287.5;  DB 2;  Length 380;
```

```
  Best Local Similarity   29.2%;  Pred. No. 1.5e-15;
  Matches   98;  Conservative   47;  Mismatches  108;  Indels   83;  Gaps   15;

Qy      33 ELKFVTLVFRHGDRGPIETFPTDPITESSW---PQGFGQLTQWGMEQHYELGSYIRKRY- 88
           :|  |  ::|||||| | ||:  |  ||: |     |:|||| || || :||  :| ||
Db      28 KLMMVQAIWRHGDRTPTETYHNDQFTENYWMFGGGGWGQLTPIGMRQHMQLGQKLRARYV 87

Qy      89 ----GRFLNDTYKHDQIYIRSTDVDRTLMSAMTNL---------------AAL------- 122
              :|||  |   :|::|||| :|||:|| :|:               |:|
Db      88 NGQPYKFLNTRYNQQEIFVRSTDKNRTLLSAFSNMVECMEIRHRKTHPLQASLKCHYVPG 147

Qy     123 --FP-----PEGISIWNPRLLWQPIPVHTVSLSEDRLL----YLPFRDCPRFEELKSETL 171
             :|     | |          : |||:||:  :|| ||    | | :|   :   |:
Db     148 VDYPDVVGWPVG---------FVPIPIHTIPDAEDHLLSVDNYCPLQD-TIWNLAKTTDT 197

Qy     172 ESEEFLKRLHPYKSFLDTLSSLSGFDDQDLFGIWSKVYDPLFCESVH------NFTLPSW 225
           |  |        | :  |::  | :: :   :|  :|: |  |  :       ||   |
Db     198 VSNYF--NSSAVTSLMGNLTNYCG-EEINPDNLWI-LYNALKIEKQYYPAQFKQFT--PW 251

Qy     226 ATEDAMIKLKELSE----------LSLLSLYGIHKQKEKSRLQGGVLVNEILKNMKLATQ 275
           |:    ::  ::           |    : ||   |   :::|| |||:|  :|   ||
Db     252 YTDSLFEQIDIVNSQVQDFQNGLGLEGKMVNGIDIGKILRKIRGGTLVNDIYNHMNRKTQ 311

Qy     276 PQKYK----------KLVMYSAHDTTVSGLQMALDV 301
                        |   |||||||:  |   | |
Db     312 CSSNNGKECTYTNRLKFFAYSAHDTTLYALFSLLGV 347


Search completed: June  4, 2002, 10:30:19
Job time: 30 sec
```

GenCore version 4.5

OM protein - protein search, using sw model

Run on: June 4, 2002, 10:30:44 ; Search time 13.4 Seconds
(without alignments)
1112.465 Million cell updates/sec

Title: US-09-402-845-2
Perfect score: 2060
Sequence: 1 MGAVPLPLSPTASLSLGFLL.....DWATECMATSSHQGTVGALG 385

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1818	88.3	381	1	PPAP_RAT	P20646 rattus norv
2	1661.5	80.7	386	1	PPAP_HUMAN	P15309 homo sapien
3	1022.5	49.6	423	1	PPAL_HUMAN	P11117 homo sapien
4	1020.5	49.5	421	1	PPAL_MOUSE	P24638 mus musculu
5	1019.5	49.5	423	1	PPAL_RAT	P20611 rattus norv
6	499.5	24.2	394	1	PPAY_CAEEL	Q10944 caenorhabdi
7	310	15.0	755	1	PPAX_CAEEL	Q09549 caenorhabdi
8	274	13.3	413	1	PPAW_CAEEL	Q09451 caenorhabdi
9	247.5	12.0	471	1	PPAV_CAEEL	Q09448 caenorhabdi
10	121	5.9	413	1	AGP_ECOLI	P19926 escherichia
11	115.5	5.6	737	1	VE02_VACCC	P21080 vaccinia vi
12	110.5	5.4	737	1	VE02_VACCV	P21604 vaccinia vi
13	110	5.3	599	1	RRPO_PEAMV	P29154 pea enation
14	109.5	5.3	737	1	VE02_VARV	P33862 variola vir
15	106	5.1	463	1	PPA2_SCHPO	Q01682 schizosacch
16	105	5.1	417	1	AGP_PRORE	Q52309 providencia
17	104.5	5.1	432	1	PPA_ECOLI	P07102 escherichia
18	104	5.0	463	1	PHYB_EMENI	O00093 emericella
19	104	5.0	489	1	PT94_YEAST	P07390 saccharomyc
20	102.5	5.0	479	1	PHYB_ASPAW	P34755 aspergillus
21	102	5.0	413	1	AGP_SALTY	O33921 salmonella
22	101	4.9	801	1	SUS2_DAUCA	O49845 daucus caro
23	99.5	4.8	937	1	MSH2_ARATH	O24617 arabidopsis
24	97.5	4.7	320	1	HLP1_HUMAN	P54257 homo sapien
25	97.5	4.7	453	1	PPA1_SCHPO	P08091 schizosacch
26	97.5	4.7	467	1	PHYA_ASPNG	P34752 aspergillus
27	97	4.7	290	1	Y290_LAMBD	P03766 bacteriopha
28	97	4.7	805	1	SUSY_PHAAU	Q01390 phaseolus a
29	97	4.7	805	1	SUSY_SOYBN	P13708 glycine max
30	96.5	4.7	503	1	CP39_RAT	P51538 rattus norv
31	96	4.7	653	1	MALQ_PYRKO	O32450 pyrococcus
32	95	4.6	820	1	SUS2_TULGE	Q41607 tulipa gesn
33	94	4.6	468	1	PPA1_PICPA	P52291 pichia past
34	94	4.6	728	1	Y376_METJA	Q57821 methanococc
35	94	4.6	806	1	SUSY_VICFA	P31926 vicia faba
36	93.5	4.5	1103	1	CHS6_USTMA	O13395 ustilago ma
37	93.5	4.5	1183	1	DRPL_RAT	P54258 rattus norv
38	93	4.5	465	1	TPSN_MOUSE	Q9r233 mus musculu
39	93	4.5	805	1	SUSY_MEDSA	O65026 medicago sa
40	92	4.5	655	1	AMYA_PYRAB	Q9v298 pyrococcus
41	91.5	4.4	452	1	HOS2_YEAST	P53096 saccharomyc
42	91.5	4.4	467	1	PHYA_ASPAW	P34753 aspergillus
43	91.5	4.4	1185	1	DRPL_HUMAN	P54259 homo sapien
44	91.5	4.4	2717	1	ZEP1_HUMAN	P15822 homo sapien
45	91	4.4	536	1	DIT1_YEAST	P21623 saccharomyc

ALIGNMENTS

```
RESULT   1
PPAP_RAT
ID   PPAP_RAT       STANDARD;      PRT;   381 AA.
AC   P20646;
DT   01-FEB-1991 (Rel. 17, Created)
DT   01-FEB-1991 (Rel. 17, Last sequence update)
DT   01-NOV-1995 (Rel. 32, Last annotation update)
DE   Prostatic acid phosphatase precursor (EC 3.1.3.2).
GN   ACPP.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=90323620; PubMed=2373368;
RA   Roiko K., Jaenne O.A., Vihko P.;
RT   "Primary structure of rat secretory acid phosphatase and comparison
RT   to other acid phosphatases.";
RL   Gene 89:223-229(1990).
RN   [2]
RP   X-RAY CRYSTALLOGRAPHY (3.0 ANGSTROMS).
RX   MEDLINE=93327749; PubMed=8334986;
RA   Schneider G., Lindqvist Y., Vihko P.;
RT   "Three-dimensional structure of rat acid phosphatase.";
RL   EMBO J. 12:2609-2615(1993).
RN   [3]
RP   X-RAY CRYSTALLOGRAPHY (3.0 ANGSTROMS).
RX   MEDLINE=94012606; PubMed=8407898;
RA   Lindqvist Y., Schneider G., Vihko P.;
RT   "Three-dimensional structure of rat acid phosphatase in complex with
RT   L(+)-tartrate.";
RL   J. Biol. Chem. 268:20744-20746(1993).
CC   -!- CATALYTIC ACTIVITY: An orthophosphoric monoester + H(2)O = an
CC     alcohol + phosphate.
CC   -!- SUBUNIT: HOMODIMER.
CC   -!- SIMILARITY: BELONGS TO THE HISTIDINE ACID PHOSPHATASE FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M32397; AAA41806.1; -.
DR   PIR; JH0152; JH0152.
DR   PDB; 1RPA; 31-MAY-94.
DR   PDB; 1RPT; 31-MAY-94.
DR   InterPro; IPR000560; His_acid_phosphtse.
DR   Pfam; PF00328; acid_phosphat; 1.
DR   PROSITE; PS00778; HIS_ACID_PHOSPHAT_2; FALSE_NEG.
DR   PROSITE; PS00616; HIS_ACID_PHOSPHAT_1; 1.
KW   Hydrolase; Glycoprotein; Signal; 3D-structure.
FT   SIGNAL       1     31
```

```
FT   CHAIN        32    381       PROSTATIC ACID PHOSPHATASE.
FT   DISULFID    160    371
FT   DISULFID    346    350
FT   ACT_SITE     43     43
FT   ACT_SITE     85     85       BY SIMILARITY.
FT   CARBOHYD     93     93       N-LINKED (GLCNAC. . .).
FT   CARBOHYD    219    219       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    332    332       N-LINKED (GLCNAC. . .).
SQ   SEQUENCE   381 AA;  43850 MW;  5EEBFF67B062FF76 CRC64;


  Query Match             88.3%;  Score 1818;  DB 1;  Length 381;
  Best Local Similarity   88.5%;  Pred. No. 3.7e-137;
  Matches  337;  Conservative   24;  Mismatches   20;  Indels    0;  Gaps    0;

Qy       1 MGAVPLPLSPTASLSLGFLLLLSLCLDPGQAKELKFVTLVFRHGDRGPIETFPTDPITES 60
           | ||||  |  ||||:||||||||| ||||||||||||||||||||||||||||| ||| ||
Db       1 MRAVPLHLVGTASLTLGFLLLLSLRLDPGQAKELKFVTLVFRHGDRGPIETFPNDPIKES 60

Qy      61 SWPQGFGQLTQWGMEQHYELGSYIRKRYGRFLNDTYKHDQIYIRSTDVDRTLMSAMTNLA 120
           ||||||||||:||| ||||||||||:|||||||::|||||:|||||||||||||||||||
Db      61 SWPQGFGQLTKWGMGQHYELGSYIRRRYGRFLNNSYKHDQVYIRSTDVDRTLMSAMTNLA 120

Qy     121 ALFPPEGISIWNPRLLWQPIPVHTVSLSEDRLLYLPFRDCPRFEELKSETLESEEFLKRL 180
           |||||||||||||||||||||||||||||||||||||||||||:|||||||:||||||||
Db     121 ALFPPEGISIWNPRLLWQPIPVHTVSLSEDRLLYLPFRDCPRFQELKSETLKSEEFLKRL 180

Qy     181 HPYKSFLDTLSSLSGFDDQDLFGIWSKVYDPLFCESVHNFTLPSWATEDAMIKLKELSEL 240
           |||||:||| |||||:||||| |||::||||:||||||||  :||||||| ||||||||
Db     181 QPYKSFIDTLPSLSGFEDQDLFEIWSRLYDPLYCESVHNFTFRTWATEDAMTKLKELSEL 240

Qy     241 SLLSLYGIHKQKEKSRLQGGVLVNEILKNMKLATQPQKYKKLVMYSAHDTTVSGLQMALD 300
           |||||||||||||||||||||||||||||||||||| :||:||||:||||||||||||:
Db     241 SLLSLYGIHKQKEKSRLQGGVLVNEILKNMKLATQPQKARKLIMYSAYDTTVSGLQMALE 300

Qy     301 VYNGVLPPYASCHMMELYHDKGGHFVEMYYRNETQNEPYPLTLPGCTHSCPLEKFAELLD 360
           .  :|||:||||||||:|||| | || |||||||||||||||||||||||||||||||||||
Db     301 LYNGLLPPYASCHIMELYQDNGGTFVEMYYRNETQNEPYPLTLPGCTHSCPLEKFAELLD 360

Qy     361 PVIPQDWATECMATSSHQGTV 381
           |||||||||||| ||:|| ::
Db     361 PVIPQDWATECMGTSNHQASL 381


RESULT   2
PPAP_HUMAN
ID   PPAP_HUMAN     STANDARD;      PRT;   386 AA.
AC   P15309;
DT   01-APR-1990 (Rel. 14, Created)
DT   01-MAR-1992 (Rel. 21, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Prostatic acid phosphatase precursor (EC 3.1.3.2).
GN   ACPP.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=92272747; PubMed=1375464;
RA   Sharief F.S., Li S.S.-L.;
RT   "Structure of human prostatic acid phosphatase gene.";
RL   Biochem. Biophys. Res. Commun. 184:1468-1476(1992).
RN   [2]
RP   SEQUENCE FROM N.A., PARTIAL SEQUENCE, DISULFIDE BONDS, AND ACTIVE
RP   SITE.
RX   MEDLINE=91115848; PubMed=1989985;
RA   van Etten R.L., Davidson R., Stevis P.E., Macarthur H., Moore D.L.;
RT   "Covalent structure, disulfide bonding, and identification of
RT   reactive surface and active site residues of human prostatic acid
RT   phosphatase.";
RL   J. Biol. Chem. 266:2313-2319(1991).
RN   [3]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=89228054; PubMed=2712834;
RA   Sharief F.S., Lee H., Leuderman M.M., Lundwall A., Deaven L.L.,
RA   Lee C.-L., Li S.S.-L.;
RT   "Human prostatic acid phosphatase: cDNA cloning, gene mapping and
RT   protein sequence homology with lysosomal acid phosphatase.";
RL   Biochem. Biophys. Res. Commun. 160:79-86(1989).
RN   [4]
RP   SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC   TISSUE=Prostate;
RX   MEDLINE=88312981; PubMed=2842184;
RA   Vihko P., Virkkunen P., Henttu P., Roiko K., Solin T., Huhtala M.L.;
RT   "Molecular cloning and sequence analysis of cDNA encoding human
RT   prostatic acid phosphatase.";
RL   FEBS Lett. 236:275-281(1988).
RN   [5]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Prostate;
RX   MEDLINE=90370491; PubMed=2395659;
RA   Tailor P.G., Govindan M.V., Patel P.C.;
RT   "Nucleotide sequence of human prostatic acid phosphatase determined
RT   from a full-length cDNA clone.";
RL   Nucleic Acids Res. 18:4928-4928(1990).
RN   [6]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=95038536; PubMed=7951074;
RA   Sharief F.S., Li S.S.-L.;
RT   "Nucleotide sequence of human prostatic acid phosphatase ACPP gene,
RT   including seven Alu repeats.";
RL   Biochem. Mol. Biol. Int. 33:561-565(1994).
RN   [7]
RP   X-RAY CRYSTALLOGRAPHY (2.9 ANGSTROMS).
RX   MEDLINE=99023966; PubMed=9804805;
RA   Lacount M.W., Handy G., Lebioda L.;
RT   "Structural origins of L(+)-tartrate inhibition of human prostatic
RT   acid phosphatase.";
RL   J. Biol. Chem. 273:30406-30409(1998).
CC   -!- CATALYTIC ACTIVITY: An orthophosphoric monoester + H(2)O = an
CC       alcohol + phosphate.
CC   -!- SUBUNIT: HOMODIMER.
CC   -!- SIMILARITY: BELONGS TO THE HISTIDINE ACID PHOSPHATASE FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M97589; AAA60021.1; -.
DR   EMBL; M97580; AAA60021.1; JOINED.
DR   EMBL; M97581; AAA60021.1; JOINED.
DR   EMBL; M97582; AAA60021.1; JOINED.
DR   EMBL; M97583; AAA60021.1; JOINED.
DR   EMBL; M97584; AAA60021.1; JOINED.
DR   EMBL; M97585; AAA60021.1; JOINED.
DR   EMBL; M97586; AAA60021.1; JOINED.
DR   EMBL; M97587; AAA60021.1; JOINED.
DR   EMBL; M97588; AAA60021.1; JOINED.
DR   EMBL; M34840; AAA69694.1; -.
DR   EMBL; M24902; AAA60022.1; -.
DR   EMBL; X52174; CAA36422.1; -.
DR   EMBL; X53605; CAA37673.1; -.
DR   EMBL; U07097; AAB60640.1; -.
DR   EMBL; U07083; AAB60640.1; JOINED.
DR   EMBL; U07085; AAB60640.1; JOINED.
DR   EMBL; U07086; AAB60640.1; JOINED.
DR   EMBL; U07088; AAB60640.1; JOINED.
DR   EMBL; U07091; AAB60640.1; JOINED.
DR   EMBL; U07092; AAB60640.1; JOINED.
DR   EMBL; U07093; AAB60640.1; JOINED.
DR   EMBL; U07095; AAB60640.1; JOINED.
```

```
DR   PIR; A32419; A32419.
DR   PIR; S01331; S01331.
DR   PIR; S11147; S11147.
DR   PIR; JH0610; JH0610.
DR   PDB; 2HPA; 16-SEP-98.
DR   MIM; 171790; -.
DR   InterPro; IPR000560; His_acid_phosphtse.
DR   Pfam; PF00328; acid_phosphat; 1.
DR   PROSITE; PS00616; HIS_ACID_PHOSPHAT_1; 1.
DR   PROSITE; PS00778; HIS_ACID_PHOSPHAT_2; 1.
KW   Hydrolase; Glycoprotein; Signal; 3D-structure.
FT   SIGNAL       1     32
FT   CHAIN       33    386       PROSTATIC ACID PHOSPHATASE.
FT   DISULFID   161    372
FT   DISULFID   215    313
FT   DISULFID   347    351
FT   ACT_SITE    44     44       NUCLEOPHILIC ACCEPTOR (BY SIMILARITY).
FT   ACT_SITE    86     86       BY SIMILARITY.
FT   ACT_SITE   289    289       PROTON DONOR.
FT   CARBOHYD    94     94       N-LINKED (GLCNAC. . .).
FT   CARBOHYD   220    220       N-LINKED (GLCNAC. . .).
FT   CARBOHYD   333    333       N-LINKED (GLCNAC. . .).
FT   CONFLICT    15     24       SLGFLFLLFF -> ALASCFCFFC (IN REF. 3
FT                               AND 4).
FT   CONFLICT    15     24       SLGFLFLLFF -> AFASCFCFFC (IN REF. 5).
FT   CONFLICT    46     46       D -> H (IN REF. 5).
FT   CONFLICT    66     73       GFGQLTQL -> WIWPTHPA (IN REF. 4).
FT   CONFLICT    66     73       GFGQLTQL -> RIWPTHPA (IN REF. 4).
FT   CONFLICT    95     95       E -> D (IN REF. 3).
FT   CONFLICT   116    116       A -> R (IN REF. 3).
FT   CONFLICT   139    139       Q -> E (IN REF. 5).
FT   CONFLICT   157    157       P -> R (IN REF. 5).
FT   CONFLICT   212    212       P -> A (IN REF. 4).
FT   CONFLICT   215    215       C -> S (IN REF. 3).
FT   CONFLICT   294    294       S -> T (IN REF. 3).
FT   CONFLICT   372    372       C -> V (IN REF. 3).
FT   CONFLICT   383    383       D -> N (IN REF. 5).
SQ   SEQUENCE   386 AA;  44566 MW;  EF81E11DFAECADEA CRC64;


  Query Match           80.7%;  Score 1661.5;  DB 1;  Length 386;
  Best Local Similarity   81.6%;  Pred. No. 1e-124;
  Matches  311;  Conservative   28;  Mismatches   41;  Indels    1;  Gaps    1;

Qy       1 MGAVPLPLSPTASLSLGFLLLLSLCLDPG-QAKELKFVTLVFRHGDRGPIETFPTDPITE 59
           | | || |:  |||||||| ||   ||    ||||||||||||||||| ||:||||||| |
Db       1 MRAAPLLLARAASLSLGFLFLLFFWLDRSVLAKELKFVTLVFRHGDRSPIDTFPTDPIKE 60

Qy      60 SSWPQGFGQLTQWGMEQHYELGSYIRKRYGRFLNDTYKHDQIYIRSTDVDRTLMSAMTNL 119
           |||||||||||| |||||||| |||||| :|||::|||:|:|||||||||||||||||||
Db      61 SSWPQGFGQLTQLGMEQHYELGEYIRKRYRKFLNESYKHEQVYIRSTDVDRTLMSAMTNL 120

Qy     120 AALFPPEGISIWNPRLLWQPIPVHTVSLSEDRLLYLPFRDCPRFEELKSETLESEEFLKR 179
           ||||||||:||||| |||||||||| ||||:|||||||:||||:||:||||:|||| ||
Db     121 AALFPPEGVSIWNPILLWQPIPVHTVPLSEDQLLYLPFRNCPRFQELESETLKSEEFQKR 180

Qy     180 LHPYKSFLDTLSSLSGFDDQDLFGIWSKVYDPLFCESVHNFTLPSWATEDAMIKLKELSE 239
           ||||| |: ||  |||   |||||||||||||||:||||||||||||||| | ||:||||
Db     181 LHPYKDFIATLGKLSGLHGQDLFGIWSKVYDPLYCESVHNFTLPSWATEDTMTKLRELSE 240

Qy     240 LSLLSLYGIHKQKEKSRLQGGVLVNEILKNMKLATQPQKYKKLVMYSAHDTTVSGLQMAL 299
           ||||||||||||||||||||||||||| :|| |||   ||||:|||||||||||||||||
Db     241 LSLLSLYGIHKQKEKSRLQGGVLVNEILNHMKRATQIPSYKKLIMYSAHDTTVSGLQMAL 300

Qy     300 DVYNGVLPPYASCHMMELYHDKGGHFVEMYYRNETQNEPYPLTLPGCTHSCPLEKFAELL 359
           |||||:||||||||: ||| :|| :||||||||||:|||||  |||||: |||||:||||:
Db     301 DVYNGLLPPYASCHLTELYFEKGEYFVEMYYRNETQHEPYPLMLPGCSPSCPLERFAELV 360

Qy     360 DPVIPQDWATECMATSSHQGT 380
           |||||||:|||| |:|||||
Db     361 GPVIPQDWSTECMTTNSHQGT 381
```

```
RESULT   3
PPAL_HUMAN
ID   PPAL_HUMAN     STANDARD;      PRT;   423 AA.
AC   P11117;
DT   01-JUL-1989 (Rel. 11, Created)
DT   01-JUL-1989 (Rel. 11, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Lysosomal acid phosphatase precursor (EC 3.1.3.2) (LAP).
GN   ACP2.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC   TISSUE=Placenta;
RX   MEDLINE=89052645; PubMed=3191910;
RA   Pohlmann R., Krentler C., Schmidt B., Schroeder W., Lorkowski G.,
RA   Culley J., Mersmann G., Geier C., Waheed A., Gottschalk S.,
RA   Grzeschik K.H., Hasikik A., von Figura K.;
RT   "Human lysosomal acid phosphatase: cloning, expression and
RT   chromosomal assignment.";
RL   EMBO J. 7:2343-2350(1988).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Leukocyte;
RX   MEDLINE=89377828; PubMed=2776754;
RA   Geier C., von Figura K., Pohlmann R.;
RT   "Structure of the human lysosomal acid phosphatase gene.";
RL   Eur. J. Biochem. 183:611-616(1989).
CC   -!- CATALYTIC ACTIVITY: An orthophosphoric monoester + H(2)O = an '
CC       alcohol + phosphate.
CC   -!- SUBCELLULAR LOCATION: Lysosomal.
CC   -!- SIMILARITY: BELONGS TO THE HISTIDINE ACID PHOSPHATASE FAMILY.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; X12548; CAA31064.1; -.
DR   EMBL; X15525; CAA33542.1; -.
DR   EMBL; X15526; CAA33542.1; JOINED.
DR   EMBL; X15527; CAA33542.1; JOINED.
DR   EMBL; X15528; CAA33542.1; JOINED.
DR   EMBL; X15529; CAA33542.1; JOINED.
DR   EMBL; X15530; CAA33542.1; JOINED.
DR   EMBL; X15531; CAA33542.1; JOINED.
DR   EMBL; X15532; CAA33542.1; JOINED.
DR   EMBL; X15533; CAA33542.1; JOINED.
DR   EMBL; X15534; CAA33542.1; JOINED.
DR   EMBL; X15535; CAA33542.1; JOINED.
DR   PIR; S06167; S06167.
DR   HSSP; P20646; 1RPA.
DR   MIM; 171650; -.
DR   InterPro; IPR000560; His_acid_phosphtse.
DR   Pfam; PF00328; acid_phosphat; 1.
DR   PROSITE; PS00616; HIS_ACID_PHOSPHAT_1; 1.
DR   PROSITE; PS00778; HIS_ACID_PHOSPHAT_2; 1.
KW   Hydrolase; Signal; Glycoprotein; Lysosome.
FT   SIGNAL       1     30
FT   CHAIN       31    423       LYSOSOMAL ACID PHOSPHATASE.
FT   DISULFID   159    370       BY SIMILARITY.
FT   DISULFID   212    310       BY SIMILARITY.
FT   DISULFID   345    349       BY SIMILARITY.
FT   ACT_SITE    42     42       NUCLEOPHILIC ACCEPTOR (BY SIMILARITY).
FT   ACT_SITE    84     84       BY SIMILARITY.
FT   ACT_SITE   286    286       PROTON DONOR (BY SIMILARITY).
FT   CARBOHYD    92     92       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   133    133       N-LINKED (GLCNAC. . .) (POTENTIAL).
```

```
FT   CARBOHYD    167    167       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    177    177       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    191    191       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    267    267       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    322    322       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    331    331       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   423 AA;  48344 MW;  3431A30B83A1E2B4 CRC64;


  Query Match             49.6%;  Score 1022.5;  DB 1;  Length 423;
  Best Local Similarity   51.4%;  Pred. No. 6.7e-74;
  Matches  189;  Conservative   62;  Mismatches  112;  Indels    5;  Gaps    3;

Qy      14 LSLGFLLLLSLCLDPGQAKELKFVTLVFRHGDRGPIETFPTDPITESSWPQGFGQLTQWG 73
           | ||  |::   : | :|: |:||||::||||| |::|:| ||  |  |||||||||: |
Db      16 LLLGVNLVV---MPPTRARSLRFVTLLYRHGDRSPVKTYPKDPYQEEEWPQGFGQLTKEG 72

Qy      74 MEQHYELGSYIRKRYGRFLNDTYKHDQIYIRSTDVDRTLMSAMTNLAALFPPEGISIWNP 133
           | ||:|||  :|:||  ||| :|   ::|:|||| |||||||  ||| |||| |:  :||
Db      73 MLQHWELGQALRQRYHGFLNTSYHRQEVYVRSTDFDRTLMSAEANLAGLFPPNGMQRFNP 132

Qy     134 RLLWQPIPVHTVSLSEDRLLYLPFRDCPRFEELKSETLESEEFLKRLHPYKSFLDTLSSL 193
            : |||||||| ::|||||  |   |||:|:|::|| :: |:          ||| :::
Db     133 NISWQPIPVHTVPITEDRLLKFPLGPCPRYEQLQNETRQTPEYQNESSRNAQFLDMVANE 192

Qy     194 SGFDDQDLFGIWSKVYDPLFCESVHNFTLPSWATEDAMIKLKELSELSLLSLYGIHKQKE 253
           :|  |  |  :|: ||| ||||  |   || ||:   | :|  | : |   |:||::| |
Db     193 TGLTDLTLETVWN-VYDTLFCEQTHGLRLPPWASPQTMQRLSRLKDFSFRFLFGIYQQAE 251

Qy     254 KSRLQGGVLVNEILKNMKLATQPQKYKKLVMYSAHDTTVSGLQMALDVYNGVLPPYASCH 313
           |:|||||: :| ||: |     :  ||::|||||:  ||||||||||   ||||||
Db     252 KARLQGGVLLAQIRKNLTLMATTSQLPKLLVYSAHDTTLVALQMALDVYNGEQAPYASCH 311

Qy     314 MMELYHDKGGHF-VEMYYRNETQNEPYPLTLPGCTHSCPLEKFAELLDPVIPQDWATECM 372
           : ||| :  |:| ||||:|||:   |:||:|||| | |||: |  | :|||:|:||  ||
Db     312 IFELYQEDSGNFSVEMYFRNESDKAPWPLSLPGCPHRCPLQDFLRLTEPVVPKDWQQECQ 371

Qy     373 ATSSHQGT 380
             |   |
Db     372 LASGPADT 379


RESULT   4
PPAL_MOUSE
ID   PPAL_MOUSE     STANDARD;      PRT;   421 AA.
AC   P24638;
DT   01-MAR-1992 (Rel. 21, Created)
DT   01-MAR-1992 (Rel. 21, Last sequence update)
DT   01-NOV-1995 (Rel. 32, Last annotation update)
DE   Lysosomal acid phosphatase precursor (EC 3.1.3.2) (LAP) (Fragment).
GN   ACP2.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=91282986; PubMed=2059337;
RA   Geier C., von Figura K., Pohlmann R.;
RT   "Molecular cloning of the mouse lysosomal acid phosphatase.";
RL   Biol. Chem. Hoppe-Seyler 372:301-304(1991).
CC   -!- CATALYTIC ACTIVITY: An orthophosphoric monoester + H(2)O = an
CC       alcohol + phosphate.
CC   -!- SUBCELLULAR LOCATION: Lysosomal.
CC   -!- SIMILARITY: BELONGS TO THE HISTIDINE ACID PHOSPHATASE FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X57199; CAA40485.1; -.
DR   PIR; S14742; S14742.
DR   HSSP; P15309; 2HPA.
DR   MGD; MGI:87882; Acp2.
DR   InterPro; IPR000560; His_acid_phosphtse.
DR   Pfam; PF00328; acid_phosphat; 1.
DR   PROSITE; PS00616; HIS_ACID_PHOSPHAT_1; 1.
DR   PROSITE; PS00778; HIS_ACID_PHOSPHAT_2; 1.
KW   Hydrolase; Signal; Glycoprotein; Lysosome.
FT   NON_TER       1      1
FT   SIGNAL       <1     28
FT   CHAIN        29    421       LYSOSOMAL ACID PHOSPHATASE.
FT   DISULFID    157    368       BY SIMILARITY.
FT   DISULFID    210    308       BY SIMILARITY.
FT   DISULFID    343    347       BY SIMILARITY.
FT   ACT_SITE     40     40       NUCLEOPHILIC ACCEPTOR (BY SIMILARITY).
FT   ACT_SITE     82     82       BY SIMILARITY.
FT   ACT_SITE    284    284       PROTON DONOR (BY SIMILARITY).
FT   CARBOHYD     90     90       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    131    131       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    165    165       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    175    175       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    189    189       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    195    195       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    265    265       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    320    320       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    329    329       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    372    372       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   421 AA;  48279 MW;  C661F431239D676D CRC64;


  Query Match             49.5%;  Score 1020.5;  DB 1;  Length 421;
  Best Local Similarity   51.4%;  Pred. No. 9.6e-74;
  Matches  187;  Conservative   66;  Mismatches  106;  Indels    5;  Gaps    3;

Qy      21 LLSLCL---DPGQAKELKFVTLVFRHGDRGPIETFPTDPITESSWPQGFGQLTQWGMEQH 77
           || :||    | ||: |:||||::||||| |::|:| ||  |  |||||||||: || ||
Db      15 LLGMCLTVMPPIQARSLRFVTLLYRHGDRSPVKTYPKDPYQEEKWPQGFGQLTKEGMLQH 74

Qy      78 YELGSYIRKRYGRFLNDTYKHDQIYIRSTDVDRTLMSAMTNLAALFPPEGISIWNPRLLW 137
           :|||  :|:||  ||| :|   ::|:|||| |||||||  ||| ||||  :  ::| : |
Db      75 WELGQALRQRYHGFLNTSYHRQEVYVRSTDFDRTLMSAEANLAGLFPPNEVQHFSPNISW 134

Qy     138 QPIPVHTVSLSEDRLLYLPFRDCPRFEELKSETLESEEFLKRLHPYKSFLDTLSSLSGFD 197
           |||||||| ::|||||  |   |||:|:|::|| :: |:  |      ||: ::: :|
Db     135 QPIPVHTVPITEDRLLKFPLGPCPRYEQLQNETRQTPEYQNRSIQNAQFLNMVANETGLT 194

Qy     198 DQDLFGIWSKVYDPLFCESVHNFTLPSWATEDAMIKLKELSELSLLSLYGIHKQKEKSRL 257
           :  |  ||: ||| ||||  |   || ||:   : :| :| : | | |:|||:| :|:||
Db     195 NVTLETIWN-VYDTLFCEQTHGLLLPPWASPQTVQRLSQLKDFSFLFLFGIHEQVQKARL 253

Qy     258 QGGVLVNEILKNMKLATQPQKYKKLVMYSAHDTTVSGLQMALDVYNGVLPPYASCHMMEL 317
           |||||: :||||: |     :: ||::|||||||:  |||||:||||   ||||||: ||
Db     254 QGGVLLAQILKNLTLMATTSQFPKLLVYSAHDTTLVALQMALNVYNGKQAPYASCHIFEL 313

Qy     318 YHDKGGHF-VEMYYRNETQNEPYPLTLPGCTHSCPLEKFAELLDPVIPQDWATECMATSS 376
           | :  |:| ||||:||:::  |:|| |||| | |||: |  | :||||:||  ||   :
Db     314 YQEDNGNFSVEMYFRNDSKKAPWPLILPGCPHRCPLQDFLRLTEPVIPKDWQKECQLAND 373

Qy     377 HQGT 380
             |
Db     374 TADT 377


RESULT   5
PPAL_RAT
ID   PPAL_RAT       STANDARD;      PRT;   423 AA.
AC   P20611;
DT   01-FEB-1991 (Rel. 17, Created)
DT   01-FEB-1991 (Rel. 17, Last sequence update)
DT   01-NOV-1995 (Rel. 32, Last annotation update)
DE   Lysosomal acid phosphatase precursor (EC 3.1.3.2) (LAP).
```

```
GN   ACP2.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=89350910; PubMed=2764916;
RA   Himeno M., Fujita H., Noguchi Y., Kono A., Kato K.;
RT   "Isolation and sequencing of a cDNA clone encoding acid phosphatase
RT   in rat liver lysosomes.";
RL   Biochem. Biophys. Res. Commun. 162:1044-1053(1989).
CC   -!- CATALYTIC ACTIVITY: An orthophosphoric monoester + H(2)O = an
CC       alcohol + phosphate.
CC   -!- SUBCELLULAR LOCATION: Lysosomal.
CC   -!- SIMILARITY: BELONGS TO THE HISTIDINE ACID PHOSPHATASE FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M27893; AAA40744.1; -.
DR   PIR; A33395; A33395.
DR   HSSP; P20646; 1RPA.
DR   InterPro; IPR000560; His_acid_phosphtse.
DR   Pfam; PF00328; acid_phosphat; 1.
DR   PROSITE; PS00616; HIS_ACID_PHOSPHAT_1; 1.
DR   PROSITE; PS00778; HIS_ACID_PHOSPHAT_2; 1.
KW   Hydrolase; Signal; Glycoprotein; Lysosome.
FT   SIGNAL        1     30
FT   CHAIN        31    423       LYSOSOMAL ACID PHOSPHATASE.
FT   DISULFID    159    370       BY SIMILARITY.
FT   DISULFID    212    310       BY SIMILARITY.
FT   DISULFID    345    349       BY SIMILARITY.
FT   ACT_SITE     42     42       NUCLEOPHILIC ACCEPTOR (BY SIMILARITY).
FT   ACT_SITE     84     84       BY SIMILARITY.
FT   ACT_SITE    286    286       PROTON DONOR (BY SIMILARITY).
FT   CARBOHYD     92     92       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    133    133       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    167    167       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    177    177       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    191    191       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    197    197       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    267    267       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    322    322       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    331    331       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   423 AA;  48319 MW;  6F4C4819CADD4496 CRC64;
```

```
  Query Match             49.5%;  Score 1019.5;  DB 1;  Length 423;
  Best Local Similarity   51.9%;  Pred. No. 1.2e-73;
  Matches  189;  Conservative   61;  Mismatches  109;  Indels    5;  Gaps    3;

Qy      21 LLSLCL---DPGQAKELKFVTLVFRHGDRGPIETFPTDPITESSWPQGFGQLTQWGMEQH 77
           || :||    | ||: |:||||::||||| |:: :| ||  |  |||||||||: || ||
Db      17 LLGMCLMVMPPIQARSLRFVTLLYRHGDRSPVKAYPKDPYQEEKWPQGFGQLTKEGMLQH 76

Qy      78 YELGSYIRKRYGRFLNDTYKHDQIYIRSTDVDRTLMSAMTNLAALFPPEGISIWNPRLLW 137
           :|||  :|:||  ||| :|   ::|:|||| |||||||  ||| |||| :  :|| : |
Db      77 WELGQALRQRYHGFLNASYHRQEVYVRSTDFDRTLMSAEANLAGLFPPTEVQHFNPNISW 136

Qy     138 QPIPVHTVSLSEDRLLYLPFRDCPRFEELKSETLESEEFLKRLHPYKSFLDTLSSLSGFD 197
           |||||||| ::|||||  |   |||:|:|::|| :: |:          ||| ::: :|
Db     137 QPIPVHTVPITEDRLLKFPLGPCPRYEQLQNETRQTPEYQNMSIQNAQFLDMVANETGLM 196

Qy     198 DQDLFGIWSKVYDPLFCESVHNFTLPSWATEDAMIKLKELSELSLLSLYGIHKQKEKSRL 257
           :  |  ||: ||| ||||  |   || ||:   :  | :| : | | |:||| | :|:||
Db     197 NLTLETIWN-VYDTLFCEQTHGLLLPPWASPQTVQALSQLKDFSFLFLFGIHDQVQKARL 255

Qy     258 QGGVLVNEILKNMKLATQPQKYKKLVMYSAHDTTVSGLQMALDVYNGVLPPYASCHMMEL 317
           |||||: :|||| |     :: ||::||||||:  |||||:||||   ||||||: ||
Db     256 QGGVLLAQILKNLTLMATTSQFPKLLVYSAHDTTLVALQMALNVYNGKQAPYASCHIFEL 315

Qy     318 YHDKGGHF-VEMYYRNETQNEPYPLTLPGCTHSCPLEKFAELLDPVIPQDWATECMATSS 376
           | :  |:| ||||:||:::  |:|||||| | |||: |  | :||||:||  ||   |
Db     316 YQEDNGNFSVEMYFRNDSKKAPWPLTLPGCPHRCPLQDFLRLTEPVIPKDWQKECQLASD 375

Qy     377 HQGT 380
               |
Db     376 TADT 379
```

```
RESULT   6
PPAY_CAEEL
ID   PPAY_CAEEL     STANDARD;      PRT;   394 AA.
AC   Q10944;
DT   01-NOV-1997 (Rel. 35, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Putative acid phosphatase B0361.7 precursor (EC 3.1.3.2).
GN   B0361.7.
OS   Caenorhabditis elegans.
OC   Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC   Rhabditidae; Peloderinae; Caenorhabditis.
OX   NCBI_TaxID=6239;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BRISTOL N2;
RA   Du Z.;
RL   Submitted (FEB-1994) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   REVISIONS.
RA   Waterston R.;
RL   Submitted (MAR-2001) to the EMBL/GenBank/DDBJ databases.
CC   -!- CATALYTIC ACTIVITY: An orthophosphoric monoester + H(2)O = an
CC       alcohol + phosphate.
CC   -!- SIMILARITY: BELONGS TO THE HISTIDINE ACID PHOSPHATASE FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U00031; AAK18870.1; -.
DR   HSSP; P20646; 1RPA.
DR   WormPep; B0361.7; CE24767.
DR   InterPro; IPR000560; His_acid_phosphtse.
DR   Pfam; PF00328; acid_phosphat; 1.
DR   PROSITE; PS00616; HIS_ACID_PHOSPHAT_1; 1.
DR   PROSITE; PS00778; HIS_ACID_PHOSPHAT_2; 1.
KW   Hypothetical protein; Signal; Glycoprotein; Hydrolase.
FT   SIGNAL        1     13       POTENTIAL.
FT   CHAIN        14    394       PUTATIVE ACID PHOSPHATASE B0361.7.
FT   ACT_SITE     40     40       NUCLEOPHILIC ACCEPTOR (BY SIMILARITY).
FT   ACT_SITE     80     80       BY SIMILARITY.
FT   ACT_SITE    250    250       PROTON DONOR (BY SIMILARITY).
FT   DISULFID    152    335       BY SIMILARITY.
FT   DISULFID    205    274       BY SIMILARITY.
FT   DISULFID    310    314       BY SIMILARITY.
FT   CARBOHYD    104    104       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    284    284       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    295    295       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   394 AA;  45011 MW;  43700D7C8BF0D8BA CRC64;
```

```
  Query Match             24.2%;  Score 499.5;  DB 1;  Length 394;
  Best Local Similarity   33.2%;  Pred. No. 2.2e-32;
  Matches  130;  Conservative   54;  Mismatches  125;  Indels   83;  Gaps   14;
```

```
Qy      19 LLLLSLCLD-----------PGQAK--ELKFVTLVFRHGDRGPIETFPTDPITESSWPQ 64
           |||| | :             | ||    |::|  |:||||| | |    | ||:  ||:
Db       2 LLLLVLLIGASGINAVVYKEVPIQANTDTLEYVHTVWRHGDRTPAELLFPDDITK--WPE 59

Qy      65 GFGQLTQWGMEQHYELGSYIRKRYGRFLNDTYKHDQIYIRSTDVDRTLMSAMTNLAALFP 124
           | |:||: |  | | || ::::||| :| : :  : ||||||:| :|||||  |:| |||
Db      60 GLGELTEQGAAQQYRLGQWLKRRYGSWLGEKFNRNAIYIRSSDYNRTLMSAQANMAGLFP 119

Qy     125 PE-----GISIWNPRLLWQPIPVHTVSLSEDRLLYLPFRDCPRFEELKSETLESE----- 174
           |:     |       |:||||||||:|   |: ||         ||    | | |
Db     120 PKYPIAGG-------LMWQPIPVHTISKPTDKELY---------EEASCPTAEIEMNAQW 163

Qy     175 EFLKRLHPYKSFLDTLSSLS---GFDDQDLFGIWSKVYDPLFCE-SVHNFTLPSWATEDA 230
           :  |     | |   ||  |      : :|   | :::| ||||   |  ||
Db     164 KSTKANGIRKKFARELSFFSQKLNLPNMELKATW-RIFDNLFCELEFHTDTL-------- 214

Qy     231 MIKLKELSELSLLSLYGIHKQKEKSRLQGGVLVNEILKNM--KLATQPQKYKKLVMYSAH 288
                                  ||:|| |: ||      | :    |  |   ||||
Db     215 ----------------------RRLRGGTLLEEIFHRFSDKASGSLGKEAKFYAYSAH 250

Qy     289 DTTVSGLQMALDVYNGVLPPYASCHMMELYH-DKGGHFVEMYYRNETQ-NEPYPLTLPGC 346
           |:|:: |   | |:  : | ||:| ::|::       : ::::|||  :     ::|||
Db     251 DSTIAALLATLGVFYDIYPKYATCLLIEMHKLANETRLIRVFHKNETDIDRLIEYSIPGC 310

Qy     347 THSCPLEKFAELLDPVIPQDWATECMATSSHQ 378
              | |:|  : |    |:||  ||   :| |
Db     311 DDPCTLQKLGDDLKKYFPEDWEAECGLKTSFQ 342


RESULT   7
PPAX_CAEEL
ID   PPAX_CAEEL     STANDARD;      PRT;   755 AA.
AC   Q09549; Q17843;
DT   01-NOV-1995 (Rel. 32, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   Putative acid phosphatase F26C11.1 (EC 3.1.3.2).
GN   F26C11.1.
OS   Caenorhabditis elegans.
OC   Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC   Rhabditidae; Peloderinae; Caenorhabditis.
OX   NCBI_TaxID=6239;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BRISTOL N2;
RA   Matthews P., Lloyd C.;
RL   Submitted (NOV-1998) to the EMBL/GenBank/DDBJ databases.
CC   -!- CATALYTIC ACTIVITY: An orthophosphoric monoester + H(2)O = an
CC       alcohol + phosphate.
CC   -!- SIMILARITY: BELONGS TO THE HISTIDINE ACID PHOSPHATASE FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; Z47072; CAA87370.1; -.
DR   EMBL; Z54342; CAA87370.1; JOINED.
DR   EMBL; Z54342; CAA91156.1; -.
DR   EMBL; Z47072; CAA91156.1; JOINED.
DR   HSSP; P20646; 1RPA.
DR   WormPep; F26C11.1; CE05732.
DR   InterPro; IPR000560; His_acid_phosphtse.
DR   Pfam; PF00328; acid_phosphat; 1.
DR   PROSITE; PS00616; HIS_ACID_PHOSPHAT_1; 1.
DR   PROSITE; PS00778; HIS_ACID_PHOSPHAT_2; 1.
KW   Hypothetical protein; Hydrolase.
FT   ACT_SITE    356    356       NUCLEOPHILIC ACCEPTOR (BY SIMILARITY).
FT   ACT_SITE    399    399       BY SIMILARITY.
FT   ACT_SITE    636    636       PROTON DONOR (BY SIMILARITY).
FT   DISULFID    702    708       BY SIMILARITY.
SQ   SEQUENCE   755 AA;  87783 MW;  E7F3050115A235EB CRC64;


  Query Match             15.0%;  Score 310;  DB 1;  Length 755;
  Best Local Similarity   26.7%;  Pred. No. 5.9e-17;
  Matches  112;  Conservative   65;  Mismatches  148;  Indels   94;  Gaps   16;

Qy      33 ELKFVTLVFRHGDRGPIETFPTDPITESSWP---QGFGQLTQWGMEQHYELGSYIRKRY- 88
           :|:||  ::|||||  :|     ||:| :|     | |:||  || :   ||:  |:||
Db     346 KLEFVQTIWRHGDRSALEGL--FPISEKNWTFGGGGLGELTPMGMSEMNNLGTIFRRRYV 403

Qy      89 --GRFLNDTYKHDQIYIRSTDVDRTLMSAMTNLAALFPPEGISIWNPR-------LLWQP 139
             :||:  |   :||||||:::||::|||: |  :|||     || :       : ||
Db     404 EDQQFLSHRYAAKEIYIRSTNLNRTIISAMSLLYGMFPP---GAWNIQGVDYPNDVDWQQ 460

Qy     140 ----IPVHTVSLSEDRLLYLPFRDCPRFEELKSE------------------------- 169
               ||||   :  |:       :| ||:||: :
Db     461 GFTFIPVHVDGI--DQCAVAQLCNCRRFQELQEKWAELDEVKNATVAMIALNRRVAAFYN 518

Qy     170 -TLESEEFLKRLHPYKS----FLDTLSSLSGFDDQDLFGIWSKVYDPLFCESVHNFTLPS 224
            | : |:| :    :|     | ||:    : ::||:    : | |    :  ||  |
Db     519 VTDQPEKFNRYTDAWKCQRNWFNDTMYQQLPWYNEDLYNEAQRTYAPFKRFTEGNFGNPK 578

Qy     225 WATEDAMIKLKELSELSLLSLYGIHKQKEKSRLQGGVLVNEILKNMK-----------LA 273
           :  |                  ||   :| | |||| |:||| :  :           :
Db     579 PSIVD----------------GIDIPQEVSTLQGGPLLNEIFERGREKIRCVADAENCS 621

Qy     274 TQPQKYKKLVMYSAHDTTVSGLQMALDVYNGV-----LPPYASCHMMELYHDKGGH-FVE 327
           |  |   ||:||  |  | : | :  |      |  :|   :| | : |    |:
Db     622 IDYLKPLKFYAYSSHDQLVYALLVTLGITDVVKTVDGWPDTSSSLTIEYYSNPGNQSSVK 681

Qy     328 MYYRNETQNEPYPLT--LPGCTHS--CPLEKFAELLDPVIP-QDWATECMATSSHQGTV 381
             ||: : :    :|  :| |  :  | :  |  :      |  |: | |  : |   :|
Db     682 FLYRDNSNDNFSDVTSQIPVCNGAQYCAMSDFQNIAYQFKPLPDYMTLCETSLSSISSV 740


RESULT   8
PPAW_CAEEL
ID   PPAW_CAEEL     STANDARD;      PRT;   413 AA.
AC   Q09451;
DT   01-NOV-1997 (Rel. 35, Created)
DT   15-DEC-1998 (Rel. 37, Last sequence update)
DT   15-DEC-1998 (Rel. 37, Last annotation update)
DE   Putative acid phosphatase C05C10.4 (EC 3.1.3.2).
GN   C05C10.4.
OS   Caenorhabditis elegans.
OC   Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC   Rhabditidae; Peloderinae; Caenorhabditis.
OX   NCBI_TaxID=6239;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BRISTOL N2;
RA   Matthews P.;
RL   Submitted (FEB-1995) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   REVISIONS.
RC   STRAIN=BRISTOL N2;
RA   Jones S.J.M.;
RL   Submitted (JUL-1998) to the EMBL/GenBank/DDBJ databases.
CC   -!- CATALYTIC ACTIVITY: An orthophosphoric monoester + H(2)O = an
CC       alcohol + phosphate.
CC   -!- SIMILARITY: BELONGS TO THE HISTIDINE ACID PHOSPHATASE FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; Z48178; CAA88205.1; -.
```

```
DR   WormPep; C05C10.4; CE17370.
DR   InterPro; IPR000560; His_acid_phosphtse.
DR   Pfam; PF00328; acid_phosphat; 1.
DR   PROSITE; PS00616; HIS_ACID_PHOSPHAT_1; 1.
DR   PROSITE; PS00778; HIS_ACID_PHOSPHAT_2; 1.
KW   Hypothetical protein; Hydrolase.
FT   ACT_SITE     35     35       NUCLEOPHILIC ACCEPTOR (BY SIMILARITY).
FT   ACT_SITE     78     78       BY SIMILARITY.
FT   ACT_SITE    314    314       PROTON DONOR (BY SIMILARITY).
FT   DISULFID    381    387       BY SIMILARITY.
FT   DOMAIN       55     60       POLY-GLY.
SQ   SEQUENCE   413 AA;  46617 MW;  BFFC70995EA903F0 CRC64;


  Query Match             13.3%;  Score 274;  DB 1;  Length 413;
  Best Local Similarity   28.1%;  Pred. No. 1.9e-14;
  Matches  119;  Conservative   62;  Mismatches  153;  Indels   90;  Gaps   25;

Qy      18 FLLLLS---LCLDPGQAK-ELKFVTLVFRHGDR-GPIETFPTDPITESSW---PQGFGQL 69
           | || |   |    ||:  :|:||  ::|||:|   :: :   || |  |     | |:|
Db       6 FSLLASTAILVFANGQSNVKLEFVQAMWRHGERASQVDQY---PIYEKDWIYGGGGLGEL 62

Qy      70 TQWGMEQHYELGSYIRKRYG---RFLNDTYKHDQIYIRSTDVDRTLMSAMTNLAALFPPE 126
           |  || :  |||  |||||    :||   |   ::| |||: :||::|| : |  ||||
Db      63 TAIGMGEMNELGWLIRKRYVTKLKFLTPKYASREVYFRSTNFNRTIISAQSLLYGLFPPS 122

Qy     127 GISIWN-----PRLLWQP----IPVHTVSLSEDRLLYLPFRDCPRFEELKSETLESEEFL 177
             : |       | | |    :|||      |:        | |:: |: : |   | |
Db     123 LYDVKNVDYPYSPLTWFPGFTFVPVHVD--GPDQCAASQNCPCTRYDLLQGQMLTLPEVL 180

Qy     178 ----------KRLHPYKSFLDTLSSLSGFDDQDLFGIWSKVYDPLFCESVH-NFTLPS-- 224
                    :|:  | :    | | : : |      |         |:  : | |: :
Db     181 PKYTQVVLLNRRVGGYYNMTSGLDSFTTYPD-----TWK-------CQRAYFNRTMYAKL 228

Qy     225 -WATED---------AMIK--LKELSELSLLSLYGIHKQKEKSRLQGGVLVNEILK--NM 270
            |  |:         | :|  |:   |   ::  |:    |  ::: ||::||:    |
Db     229 PWYNEELYYQAQVTYAPVKGFLEGNFENPAVTSSGLDVGLEIKKVRSGVIINEVFNRANE 288

Qy     271 KL--------ATQPQKYKKLVMYSAHDTTVSGLQMALDV------YNGVLPPYASCHMME 316
           ||         |      |   || ||  | |: :|| :       :|  | ||:   ||
Db     289 KLNCAELGQNCTSYLNKLKFYGYSIHDNNVYGVLVALGIPQIANTLDG-WPAYAAGIFME 347

Qy     317 LYHDKGGH--FVEMYYR--NETQNEPYPLTLPGCTHS--CP---LEKFAELLDPVIPQDW 367
           : :   :  | :: ||  ::|        || | :  ||   |:  || | |:   |
Db     348 FHRNTSTNERFFKVLYREGDDTPISDVTSQLPICNGATLCPLGALQTLAETLKPL--PDI 405

Qy     368 ATEC 371
           | |
Db     406 TTLC 409


RESULT   9
PPAV_CAEEL
ID   PPAV_CAEEL     STANDARD;      PRT;   471 AA.
AC   Q09448;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   01-NOV-1997 (Rel. 35, Last annotation update)
DE   Putative acid phosphatase C05C10.1 (EC 3.1.3.2) (Fragment).
GN   C05C10.1.
OS   Caenorhabditis elegans.
OC   Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC   Rhabditidae; Peloderinae; Caenorhabditis.
OX   NCBI_TaxID=6239;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BRISTOL N2;
RA   Matthews P.;
RL   Submitted (FEB-1995) to the EMBL/GenBank/DDBJ databases.
CC   -!- CATALYTIC ACTIVITY: An orthophosphoric monoester + H(2)O = an
CC       alcohol + phosphate.
CC   -!- SIMILARITY: BELONGS TO THE HISTIDINE ACID PHOSPHATASE FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; Z48178; CAA88204.1; -.
DR   WormPep; C05C10.1; CE01465.
DR   InterPro; IPR000560; His_acid_phosphtse.
DR   Pfam; PF00328; acid_phosphat; 1.
DR   PROSITE; PS00616; HIS_ACID_PHOSPHAT_1; FALSE_NEG.
DR   PROSITE; PS00778; HIS_ACID_PHOSPHAT_2; 1.
KW   Hypothetical protein; Hydrolase.
FT   NON_TER       1      1
FT   ACT_SITE      4      4       NUCLEOPHILIC ACCEPTOR (BY SIMILARITY).
FT   ACT_SITE     47     47       BY SIMILARITY.
FT   ACT_SITE    283    283       PROTON DONOR (BY SIMILARITY).
FT   DISULFID    350    356       BY SIMILARITY.
FT   DOMAIN       24     29       POLY-GLY.
SQ   SEQUENCE   471 AA;  53648 MW;  B3D81126E1AD2A94 CRC64;


  Query Match             12.0%;  Score 247.5;  DB 1;  Length 471;
  Best Local Similarity   26.9%;  Pred. No. 2.9e-12;
  Matches  101;  Conservative   57;  Mismatches  161;  Indels   57;  Gaps   17;

Qy      40 VFRHGDRGPIETFPTDPITESSW---PQGFGQLTQWGMEQHYELGSYIRKRYGR---FLN 93
           ::|||:|  :      || |  |     | |:||  || :   ||  ||:|| |   ||
Db       1 MWRHGERSALADL--YPIYEKDWVFGGGGLGELTGRGMGEMNNLGRLIRERYVRKFNFLE 58

Qy      94 DTYKHDQIYIRSTDVDRTLMSAMTNLAALFPPEGISIWN-----PRLLWQP----IPVHT 144
             |   ::| |||:::||::|||: |  ||||    | |       | | |    :|||
Db      59 PKYASKEVYFRSTNLNRTIISAMSLLYGLFPPSLYDIPNVDYPFTPLKWLPGLAFVPVHV 118

Qy     145 VSLSEDRLLYLPFRDCPRFEELKSETLESEE----FLKRLHPYKSFLDTLSSLSGFDDQD 200
                |:       |||:: |: : |   |    | : :   :      :  :| |
Db     119 D--GSDQCAASQNCPCPRYDFLQQQMLTLPEVQQAFQQVILLNRQIAPLYNVTTGVDTFY 176

Qy     201 LFGIWSKVYDPLFCESVHNFTLPSWATEDAMIK-------LKELSELSL----LSLYGIH 249
           ::    |     | :::::  || |  |    |       :|   | |      ::  |:
Db     177 VYPDTWKCQRAYFNKTMYD-KLP-WYNEQLYSKAEITYAPIKGFLEGSFSQPAVTSNGLD 234

Qy     250 KQKEKSRLQGGVLVNEILK--NMKL--------ATQPQKYKKLVMYSAHDTTVSGLQMAL 299
             |  ::: ||::|||:   : ||          |       |   || ||  |  : :||
Db     235 VAFEIQQVRSGVMINEIVSRASEKLNCVERGQNCTSYLNKLKFYGYSIHDNNVYAVLVAL 294

Qy     300 DV-----YNGVLPPYASCHMMELYHDK--GGHFVEMYYRNETQNEPYPLT--LPGC--TH 348
           :          | ||:    | | :         :: || :  ::   :|  :| |
Db     295 GIPHISATEDGWPSYAAAIFFEFYRNSQTNKRLFKVLYRQDASSQITDVTSQVPMCQGVS 354

Qy     349 SCPLEKFAELLDPVIP 364
           |||  |  | | : |
Db     355 MCPLSTFQHLADVLKP 370


RESULT  10
AGP_ECOLI
ID   AGP_ECOLI      STANDARD;      PRT;   413 AA.
AC   P19926;
DT   01-FEB-1991 (Rel. 17, Created)
DT   01-FEB-1991 (Rel. 17, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Glucose-1-phosphatase precursor (EC 3.1.3.10) (G1Pase).
GN   AGP OR B1002.
OS   Escherichia coli.
OC   Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC   Escherichia.
OX   NCBI_TaxID=562;
RN   [1]
RP   SEQUENCE FROM N.A., AND SEQUENCE OF 23-31.
RC   STRAIN=K12;
```

```
RX   MEDLINE=90130318; PubMed=2153660;
RA   Pradel E., Marck C., Boquet P.L.;
RT   "Nucleotide sequence and transcriptional analysis of the Escherichia
RT   coli agp gene encoding periplasmic acid glucose-1-phosphatase.";
RL   J. Bacteriol. 172:802-807(1990).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=K12 / MG1655;
RX   MEDLINE=97426617; PubMed=9278503;
RA   Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA   Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA   Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA   Mau B., Shao Y.;
RT   "The complete genome sequence of Escherichia coli K-12.";
RL   Science 277:1453-1474(1997).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=K12;
RX   MEDLINE=97061202; PubMed=8905232;
RA   Oshima T., Aiba H., Baba T., Fujita K., Hayashi K., Honjo A.,
RA   Ikemoto K., Inada T., Itoh T., Kajihara M., Kanai K., Kashimoto K.,
RA   Kimura S., Kitagawa M., Makino K., Masuda S., Miki T., Mizobuchi K.,
RA   Mori H., Motomura K., Nakamura Y., Nashimoto H., Nishio Y., Saito N.,
RA   Sampei G., Seki Y., Tagami H., Takemoto K., Wada C., Yamamoto Y.,
RA   Yano M., Horiuchi T.;
RT   "A 718-kb DNA sequence of the Escherichia coli K-12 genome
RT   corresponding to the 12.7-28.0 min region on the linkage map.";
RL   DNA Res. 3:137-155(1996).
RN   [4]
RP   SEQUENCE OF 23-34.
RC   STRAIN=K12 / EMG2;
RX   MEDLINE=97443975; PubMed=9298646;
RA   Link A.J., Robison K., Church G.M.;
RT   "Comparing the predicted and observed properties of proteins encoded
RT   in the genome of Escherichia coli K-12.";
RL   Electrophoresis 18:1259-1313(1997).
CC   -!- FUNCTION: ABSOLUTELY REQUIRED FOR THE GROWTH OF E.COLI IN A HIGH-
CC       PHOSPHATE MEDIUM CONTAINING G-1-P AS THE SOLE CARBON SOURCE.
CC   -!- CATALYTIC ACTIVITY: D-glucose 1-phosphate + H(2)O = D-glucose +
CC       phosphate.
CC   -!- ENZYME REGULATION: INDEPENDENT FROM INORGANIC PHOSPHATE
CC       AVAILABILITY, AND APPARENTLY SUBMITTED TO CATABOLITE REPRESSION,
CC       IT IS POSITIVELY CONTROLLED BY CAMP AND THE CAMP RECEPTOR PROTEIN.
CC   -!- SUBUNIT: HOMODIMER.
CC   -!- SUBCELLULAR LOCATION: Periplasmic.
CC   -!- MISCELLANEOUS: OPTIMALLY ACTIVE AROUND PH 4.
CC   -!- SIMILARITY: BELONGS TO THE HISTIDINE ACID PHOSPHATASE FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M33807; AAA23426.1; -.
DR   EMBL; AE000202; AAC74087.1; -.
DR   EMBL; D90737; BAA35769.1; -.
DR   EMBL; D90738; BAA35779.1; -.
DR   PIR; JV0087; JV0087.
DR   HSSP; P07102; 1DKM.
DR   EcoGene; EG10033; agp.
DR   InterPro; IPR000560; His_acid_phosphtse.
DR   Pfam; PF00328; acid_phosphat; 1.
DR   PROSITE; PS00616; HIS_ACID_PHOSPHAT_1; 1.
DR   PROSITE; PS00778; HIS_ACID_PHOSPHAT_2; 1.
KW   Hydrolase; Periplasmic; Signal; Complete proteome.
FT   SIGNAL        1     22
FT   CHAIN        23    413       GLUCOSE-1-PHOSPHATASE.
FT   ACT_SITE     40     40       NUCLEOPHILIC ACCEPTOR (BY SIMILARITY).
FT   ACT_SITE    311    311       PROTON DONOR (BY SIMILARITY).
SQ   SEQUENCE   413 AA;  45683 MW;  ADADAD3639D0D6AB CRC64;
```

```
  Query Match             5.9%;  Score 121;  DB 1;  Length 413;
  Best Local Similarity   20.0%;  Pred. No. 0.027;
  Matches   84;  Conservative   70;  Mismatches  184;  Indels   82;  Gaps   17;

Qy      12 ASLSLGFLLLLSLCLDPGQAK------ELKFVTLVFRHGDRGPIETFPT--DPITESSWP 63
           |:   | :|| |       ||:      :|: | :: ||  | |:    :  :  | : ||
Db       7 AAAVAGIVLLAS----NAQAQTVPEGYQLQQVLMMSRHNLRAPLANNGSVLEQSTPNKWP 62

Qy      64 Q---GFGQLTQWGMEQHYELGSYIRKRYGRFLNDTYKHDQ------IYIRSTDVDRTLMS 114
           :     ||||  |       :| |:|:          |  :      :|  :  : ||: :
Db      63 EWDVPGGQLTTKGGVLEVYMGHYMREWLAE--QGMVKSGECPPPYTVYAYANSLQRTVAT 120

Qy     115 AMTNLAALFPPEGISIWNPRLLWQPIPVHTVSLSEDRLLYLPFRDCPRFEEL-KSETLES 173
           |   :   ||   | : :   :    |     :::|   :          :|| | :  :|
Db     121 AQFFITGAFPGCDIPVHHQEKMGTMDPTFNPVITDDSAAFSEQAVAAMEKELSKLQLTDS 180

Qy     174 EEFLKRLHPYK-----------SFLDTLSSLSGFDDQD-----LFGIWSKVYDPLFCESV 217
            : |:::  ||           | :|  :: |   |:          : : : |    :
Db     181 YQLLEKIVNYKDSPACKEKQQCSLVDGKNTFSAKYQQEPGVSGPLKVGNSLVDAFTLQYY 240

Qy     218 HNFTLP--SWATEDAMIKLKELSELSLLSLYGIHKQKEKSRLQGGVLVNEILKNMKLATQ 275
             | :   :|    :  : | ||:|       :    | :|     ||: | |   | |
Db     241 EGFPMDQVAWGEIKSDQQWKVLSKLKNGYQDSLFTSPEVARNVAKPLVSYIDK--ALVTD 298

Qy     276 PQKYKKLVMYSAHDTTVSGLQMALDVYNGVLPPYASCH-----------MMELYHDKGG 323
             |: :   ||: :: |  |||       ||   |              : : :||
Db     299 RTSAPKITVLVGHDSNIASLLTALD-----FKPY-QLHDQNERTPIGGKIVFQRWHDSKA 352

Qy     324 H----FVEMYYRNETQ-----------NEPYPLTLPGCTHS----CPLEKFAELLDPVI 363
           :     :|  |::  |            :   | | ||       ||::||  :|:  :
Db     353 NRDLMKIEYVYQSAEQLRNADALTLQAPAQRVTLELSGCPIDADGFCPMDKFDSVLNEAV 412
```

```
RESULT  11
VE02_VACCC
ID   VE02_VACCC     STANDARD;      PRT;   737 AA.
AC   P21080;
DT   01-FEB-1991 (Rel. 17, Created)
DT   01-FEB-1991 (Rel. 17, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Protein E2.
GN   E2L.
OS   Vaccinia virus (strain Copenhagen).
OC   Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
OC   Orthopoxvirus.
OX   NCBI_TaxID=10249;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=91021027; PubMed=2219722;
RA   Goebel S.J., Johnson G.P., Perkus M.E., Davis S.W., Winslow J.P.,
RA   Paoletti E.;
RT   "The complete DNA sequence of vaccinia virus.";
RL   Virology 179:247-266(1990).
RN   [2]
RP   COMPLETE GENOME.
RA   Goebel S.J., Johnson G.P., Perkus M.E., Davis S.W., Winslow J.P.,
RA   Paoletti E.;
RT   "Appendix to 'The complete DNA sequence of vaccinia virus'.";
RL   Virology 179:517-563(1990).
CC   -!- SIMILARITY: TO VACCINIA VIRUS PROTEIN E2.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M35027; AAA48039.1; -.
DR   PIR; F42508; F42508.
```

```
SQ   SEQUENCE   737 AA;  85916 MW;  4D8EB8A1BFC2E2F4 CRC64;

  Query Match             5.6%;  Score 115.5;  DB 1;  Length 737;
  Best Local Similarity   21.5%;  Pred. No. 0.16;
  Matches   83;  Conservative   67;  Mismatches  125;  Indels  111;  Gaps   21;

Qy      10 PTASLSLGFLLLLSLCLDPGQAKELKFVTLVFRHGDRGPIETF---PTDPITESSWPQGF 66
           |  || :   |:: :|:      ::|:| :|     |   |:|      ||||| ::
Db     350 PVTSLPIHSTLVMVMCI------QMKYVDIV-EFLDEIDIDTLIEKGADPITEYTF---- 398

Qy      67 GQLTQWGMEQHYELGSYIRKRYG-------RFLND---TYKHDQIYIRSTDVDRTLMSAM 116
             |:|   :| :| :   |:||       | : :   | :     ::: | :|
Db     399 --TTRW-YNKHNDLITLYIKKYGFCPMMMKRLMFEYPLTKEASDHLLKTMDENRG----- 450

Qy     117 TNLAALFPPEGIS------IWNPRLLWQPIPV-----HTVSLSEDRLLYLPFRDCPRFEE 165
           | :| |  |         | :|: :|||      : |    :|:|      |  |:
Db     451 ---AIMFFPRTICTLPYLLCCNYKLIQKPIPFKEENRNIVYKKNNRVL------C--FDS 499

Qy     166 LKSETLESEEFLKRLHPYKSFLDTLSSLSGFDDQDLFGIWSKVYDPLFC------ESVHN 219
           |::   :|          |  : |: |    ::  |  :  : : |      ||:||
Db     500 LENSAFKS------------LIKIDSIPGLKTYNMKDITYEKSNNIICVRFIPQESIHN 546

Qy     220 FTLPSWATEDAMIKLKELSELSLLSLYGIHKQKEKSRLQGGVLVNEILKNMKLATQPQKY 279
                 |:  ||| :| ::: |: ||::    :       :|| |    :  | |||
Db     547 --------EERRIKL-QLFDIARLASYGLYYIPSRYLSSWTPVVNMI--EGREYTNPQKI 595

Qy     280 KKLV---MYSAH----------------DTTVSGLQMALDVYNGVLPPY---ASCHMME 316
           : ||   ::|                  : |:|  | |::     |  |    |   :
Db     596 ECLVILDLFSEEFIEYQNLGNAVSNKYELEYTISNYQAAINCLMSTLLIYLVLGSIRSIS 655

Qy     317 LYHDKGGHFVEMYYR----NETQNEP 338
           .    |     : ::|:    ||  :||
Db     656 KTEDFVLSILNIFYKGLKINELLSEP 681


RESULT  12
VE02_VACCV
ID   VE02_VACCV     STANDARD;      PRT;   737 AA.
AC   P21604;
DT   01-MAY-1991 (Rel. 18, Created)
DT   01-NOV-1991 (Rel. 20, Last sequence update)
DT   01-NOV-1991 (Rel. 20, Last annotation update)
DE   Protein E2.
GN   E2L.
OS   Vaccinia virus (strain WR).
OC   Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
OC   Orthopoxvirus.
OX   NCBI_TaxID=10254;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Gershon P.D., Jones E.V., Moss B., Ahn B.-Y.;
RL   Submitted (JUL-1990) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE OF 1-597 FROM N.A.
RX   MEDLINE=90377234; PubMed=2398897;
RA   Ahn B.-Y., Gershon P.D., Jones E.V., Moss B.;
RT   "Identification of rpo30, a vaccinia virus RNA polymerase gene with
RT   structural similarity to a eucaryotic transcription elongation
RT   factor.";
RL   Mol. Cell. Biol. 10:5433-5441(1990).
CC   -!- SIMILARITY: TO VACCINIA VIRUS PROTEIN E2.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M36339; AAB59822.1; -.
DR   PIR; A35928; A35928.
```

```
SQ   SEQUENCE   737 AA;  85957 MW;  D5BAF09E2F944A9E CRC64;

  Query Match             5.4%;  Score 110.5;  DB 1;  Length 737;
  Best Local Similarity   21.2%;  Pred. No. 0.4;
  Matches   82;  Conservative   68;  Mismatches  125;  Indels  111;  Gaps   21;

Qy      10 PTASLSLGFLLLLSLCLDPGQAKELKFVTLVFRHGDRGPIETF---PTDPITESSWPQGF 66
           |  || :   |:: :|:      ::|:| :|     |   |:|      ||||| ::
Db     350 PVTSLPIHSTLVMVMCI------QMKYVDIV-EFLDEIDIDTLIEKGADPITEYTF---- 398

Qy      67 GQLTQWGMEQHYELGSYIRKRYG-------RFLND---TYKHDQIYIRSTDVDRTLMSAM 116
             |:|   :| :| :   |:||       | : :   | :     ::: | :|
Db     399 --TTRW-YNKHNDLITLYIKKYGFCPMMMKRLMFEYPLTKEASDHLLKTMDENRG----- 450

Qy     117 TNLAALFPPEGIS------IWNPRLLWQPIPV-----HTVSLSEDRLLYLPFRDCPRFEE 165
           | :| |  |         | :|: :|||      : |    :|:|      |  |:
Db     451 ---AIMFFPRTICTLPYLLCCNYKLIQKPIPFKEENRNIVYKKNNRVL------C--FDS 499

Qy     166 LKSETLESEEFLKRLHPYKSFLDTLSSLSGFDDQDLFGIWSKVYDPLFC------ESVHN 219
           }::   :|          |  : |: |    ::  |  :  : : |      ||:||
Db     500 LENSAFKS------------LIKIDSIPGLKTYNMKDITYEKSNNIICVRFIPQESIHN 546

Qy     220 FTLPSWATEDAMIKLKELSELSLLSLYGIHKQKEKSRLQGGVLVNEILKNMKLATQPQKY 279
                 |:  ||| :| ::: |: ||::    :       :|| |    :  | |||
Db     547 --------EERRIKL-QLFDIARLASYGLYYIPSRYLSSWTPVVNMI--EGREYTNPQKI 595

Qy     280 KKLV---MYSAH----------------DTTVSGLQMALDVYNGVLPPY---ASCHMME 316
           : ||   ::|                  : |:|  | |::     |  |    |   :
Db     596 ECLVILDLFSEEFIEYQNLGNAVSNKYELEYTISNYQAAINCLMSTLLIYLVLGSIRSIS 655

Qy     317 LYHDKGGHFVEMYYR----NETQNEP 338
           :    : ::|:    ||  :||
Db     656 RTENFVLSILNIFYKGLKINELLSEP 681


RESULT  13
RRPO_PEAMV
ID   RRPO_PEAMV     STANDARD;      PRT;   599 AA.
AC   P29154;
DT   01-DEC-1992 (Rel. 24, Created)
DT   01-DEC-1992 (Rel. 24, Last sequence update)
DT   15-DEC-1998 (Rel. 37, Last annotation update)
DE   RNA-directed RNA polymerase (EC 2.7.7.48).
OS   Pea enation mosaic virus (PEMV).
OC   Viruses; ssRNA positive-strand viruses, no DNA stage; Luteoviridae;
OC   Enamovirus.
OX   NCBI_TaxID=12290;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=WSG;
RX   MEDLINE=91341468; PubMed=1875194;
RA   Demler S.A., de Zoeten G.A.;
RT   "The nucleotide sequence and luteovirus-like nature of RNA 1 of an
RT   aphid non-transmissible strain of pea enation mosaic virus.";
RL   J. Gen. Virol. 72:1819-1834(1991).
CC   -!- SIMILARITY: TO LUTEOVIRUS RNA-DIRECTED RNA POLYMERASES.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; L04573; AAA72297.1; ALT_INIT.
DR   PIR; JQ1384; RRXBPM.
DR   InterPro; IPR001795; Luteo_RNA_pol.
DR   Pfam; PF02123; Luteo_ORF3; 1.
DR   PRINTS; PR00914; LVIRUSRNAPOL.
KW   Transferase; RNA-directed RNA polymerase.
FT   VARIANT      66     66       G -> S.
```

```
FT   VARIANT      78     78       Q -> P.
FT   VARIANT      84     84       S -> I.
FT   VARIANT      93     93       T -> I.
FT   VARIANT      95     95       R -> G.
FT   VARIANT     104    104       R -> H.
FT   VARIANT     336    336       V -> L.
FT   VARIANT     359    359       I -> K.
FT   VARIANT     410    410       S -> F.
SQ   SEQUENCE   599 AA;  67222 MW;  7E7613477FA6E719 CRC64;


  Query Match             5.3%;  Score 110;  DB 1;  Length 599;
  Best Local Similarity   22.0%;  Pred. No. 0.33;
  Matches   73;  Conservative   59;  Mismatches  142;  Indels   58;  Gaps   15;

Qy      75 EQHYELGSYIRKRYGRFLNDTYKHDQIYIRSTDVDRTLMSAMTNLAALF----PPEGISI 130
           |    :|   :| |  | ||  :          |  : :   : :   ||    | :   |
Db     272 ESTEDLWETVRDRLFRLLNQDF---------IDPVQAVKDGLVDPIRLFVKLEPHKMEKI 322

Qy     131 WNPRLLWQPIPVHTVSLSEDRLLYLPFRDCPRFEELKSETLESEEFL--KRLHPYKSFLD 188
           | |       : :||: :  :  : |||    | |:   : |:  |   : |   :| :
Db     323 RNKRYRL----IASVSIVDQLVARMLFRDQNEEELLQHMAIPSKPGLGFSQDHQVLAFTE 378

Qy     189 TLSSLSGFDDQDLFGIWSKVYDPLFCESVHNFTLPSWATEDAMIKLKELSELSLLSLYGI 248
           ::::|:|   |||   ||:   |  | |  ::::| |  ||    |   :||:|   :|:
Db     379 SVAALAGTSAQDLVDNWSRYLTPTDC-SGFDWSVPMWLLED---DLAVRNELTLGLPHGL 434

Qy     249 HKQKEK---------SRLQGGVLVNEILKNMK-------LATQPQKYKKLVMYSAHDTTV 292
           | :|           |  |:|: :    ::        :|  :    | :|:     |
Db     435 RKMRETWLKCLGQSVFCLSNGLLLAQTSPGIQKSGSFNTSSTNSRMRYMLALYAGASWAV 494

Qy     293 SGLQMALDVYNGVLPPYA----SCHMMELYHDKGGHFV---EMYYRNETQNEPYPLTLPG 345
           :    ||:     |  ||     |   | : |   |      ::       :   | | | |
Db     495 TMGDDALESVGSDLSQYARLGIKCERAEEF-DFCSHLFRAPDVVIPKNLEKMVYGL-LSG 552

Qy     346 CTHSCPL--EKFA------ELLDPV--IPQDW 367
           .   :   ||  ::|:       :|: :  :|||:
Db     553 TSPESPLLADRFSWLSALQSILEEMRHMPQDF 584


RESULT  14
VE02_VARV
ID   VE02_VARV      STANDARD;      PRT;   737 AA.
AC   P33862;
DT   01-FEB-1994 (Rel. 28, Created)
DT   01-FEB-1994 (Rel. 28, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Protein E2.
GN   E2L.
OS   Variola virus.
OC   Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
OC   Orthopoxvirus.
OX   NCBI_TaxID=10255;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=INDIA-1967 / ISOLATE IND3;
RX   MEDLINE=94152154; PubMed=8109158;
RA   Shchelkunov S.N., Blinov V.M., Resenchuk S.M., Totmenin A.V.,
RA   Sandakhchiev L.S.;
RT   "Analysis of the nucleotide sequence of a 43 kbp segment of the
RT   genome of variola virus India-1967 strain.";
RL   Virus Res. 30:239-258(1993).
RN   [2]
RP   COMPLETE GENOME.
RC   STRAIN=INDIA-1967 / ISOLATE IND3;
RX   MEDLINE=93202281; PubMed=8384129;
RA   Shchelkunov S.N., Blinov V.M., Sandakhchiev L.S.;
RT   "Genes of variola and vaccinia viruses necessary to overcome the host
RT   protective mechanisms.";
RL   FEBS Lett. 319:80-83(1993).
CC   -!- SIMILARITY: TO VACCINIA VIRUS PROTEIN E2.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X69198; CAA48984.1; -.
DR   PIR; E36841; E36841.
SQ   SEQUENCE   737 AA;  85957 MW;  F8CA3A5A6EFA17E0 CRC64;


  Query Match             5.3%;  Score 109.5;  DB 1;  Length 737;
  Best Local Similarity   21.7%;  Pred. No. 0.48;
  Matches   83;  Conservative   67;  Mismatches  128;  Indels  105;  Gaps   21;

Qy      10 PTASLSLGFLLLLSLCLDPGQAKELKFVTLVFRHGDRGPIETF---PTDPITESSWPQGF 66
           |  || :   |:: :|:      ::|:| :|    |   |:|      |||:| ::
Db     350 PVTSLPIHSTLVMVMCI------QMKYVDIV-EFLDEIDIDTLIEKGADPITEYTF---- 398

Qy      67 GQLTQWGMEQHYELGSYIRKRYG-------RFLND---TYKHDQIYIRSTDVDRTLMSAM 116
           |:|   :| :|: :   |:||       | : :   | :     ::: | :|
Db     399 --TTRW-YNKHNDLITLYIKKYGFCPMMMKRLMFEYPLTKEASDHLLKTMDENRG----- 450

Qy     117 TNLAALFPPEGIS------IWNPRLLWQPIPVHTVSLSEDR-LLYLPFRDCPRFEELKSE 169
           | :| |  |          | :|: :|||       |:| ::|       |: |::
Db     451 ---AIMFFPRTICTLPYLLCCNYKLIQKPIPFK----EENRNIVYKKTNRVLCFDLLENS 503

Qy     170 TLESEEFLKRLHPYKSFLDTLSSLSGFDDQDLFGIWSKVYDPLFC------ESVHNFTLP 223
             :|             |  : |: |    ::  |  :  : : |      ||:||
Db     504 AFKS------------LIKIDSIPGLKTYNMKDITYEKSNNIICVRFIPQESIHN---- 546

Qy     224 SWATEDAMIKLKELSELSLLSLYGIHKQKEKSRLQGGVLVNEILKNMKLATQPQKYKKLV 283
           |:  ||| :| ::: |: ||::    :       :|| |    :  | ||| : || '
Db     547 ----EERRIKL-QLFDIARLASYGLYYIPSRYLSSWTPVVNMI--EGREYTNPQKIECLV 599

Qy     284 ---MYSAH----------------DTTVSGLQMALDVYNGVLPPYASCHMMELYHDKGG 323
              ::|                  : |:|  | |::     |  |     :     |
Db     600 ILDLFSEEFIEYQNLGNAVSNKYELEYTISNYQAAINCLMSTLLIYLVLGSIRSI-SKTE 658

Qy     324 HFV----EMYYR----NETQNEP 338
           :||     ::|:    ||  :||
Db     659 NFVLSILNIFYKGLKINELLSEP 681


RESULT  15
PPA2_SCHPO
ID   PPA2_SCHPO     STANDARD;      PRT;   463 AA.
AC   Q01682;
DT   01-JUL-1993 (Rel. 26, Created)
DT   01-JUL-1993 (Rel. 26, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Thiamine-repressible acid phosphatase precursor (EC 3.1.3.2).
GN   PHO4 OR SPBC428.03C.
OS   Schizosaccharomyces pombe (Fission yeast).
OC   Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC   Schizosaccharomycetales; Schizosaccharomycetaceae;
OC   Schizosaccharomyces.
OX   NCBI_TaxID=4896;
RN   [1]
RP   SEQUENCE FROM N.A., AND SEQUENCE OF 19-30.
RX   MEDLINE=91064763; PubMed=2249257;
RA   Yang J., Schweingruber M.E.;
RT   "The structural gene coding for thiamin-repressible acid phosphatase
RT   in Schizosaccharomyces pombe.";
RL   Curr. Genet. 18:269-272(1990).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=972;
RA   Lyne M., Rajandream M.A., Barrell B.G., Baker S., Mungall K.;
RL   Submitted (NOV-1998) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: MAY DEPHOSPHORYLATE THIAMINE PHOSPHATES.
CC   -!- CATALYTIC ACTIVITY: An orthophosphoric monoester + H(2)O = an
```

```
CC       alcohol + phosphate.
CC   -!- SUBCELLULAR LOCATION: Cell wall.
CC   -!- INDUCTION: REPRESSED BY THIAMINE.
CC   -!- SIMILARITY: BELONGS TO THE HISTIDINE ACID PHOSPHATASE FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X56939; CAA40258.1; -.
DR   EMBL; AL034382; CAA22278.1; -.
DR   PIR; S14119; S14119.
DR   HSSP; P34755; 1QFX.
DR   InterPro; IPR000560; His_acid_phosphtse.
DR   Pfam; PF00328; acid_phosphat; 1.
DR   PROSITE; PS00616; HIS_ACID_PHOSPHAT_1; 1.
DR   PROSITE; PS00778; HIS_ACID_PHOSPHAT_2; 1.
KW   Hydrolase; Glycoprotein; Cell wall; Signal.
FT   SIGNAL        1     18
FT   CHAIN        19    463       THIAMINE-REPRESSIBLE ACID PHOSPHATASE.
FT   ACT_SITE     69     69       NUCLEOPHILIC ACCEPTOR (BY SIMILARITY).
FT   ACT_SITE    340    340       PROTON DONOR (BY SIMILARITY).
FT   CARBOHYD     98     98       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    104    104       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    186    186       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    221    221       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    251    251       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    328    328       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    433    433       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    439    439       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    458    458       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   463 AA;  52118 MW;  F48EAFF8BB6B234A CRC64;

  Query Match             5.1%;  Score 106;  DB 1;  Length 463;
  Best Local Similarity   20.5%;  Pred. No. 0.49;
  Matches   79;  Conservative   56;  Mismatches  142;  Indels  108;  Gaps   17;

Qy      34 LKFVTLVFRHGDRGPIE----------------------------TFPTDPI------T 58
           :| | |: ||| | |                              ::| :|:      |
Db      60 IKQVHLLQRHGSRNPTGDDTATDVSSAQYIDIFQNKLLNGSIPVNFSYPENPLYFVKHWT 119

Qy      59 ESSWPQGFGQLTQWGMEQHYELGSYIRKRYGRFLNDTYKHDQIYIRSTDVDRTLMSAMTN 118
                :   ||:  |  : ::||  : :||   | ||  :|   | :   :| : ||
Db     120 PVIKAENADQLSSSGRIELFDLGRQVFERYYE-LFDTDVYD---INTAAQERVVDSAEWF 175

Qy     119 LAALFPPEGISIWNPRLLWQPIPVHTVSLSED--------RLLYLPFRDCPRFEE--LKS 168
              :|   |  : |          : : | ||         : |    || :|:  :
Db     176 SYGMF---GDDMQN--------KTNFIVLPEDDSAGANSLAMYY----SCPVYEDNNIDE 220

Qy     169 ETLESEE------FLK----RLHPYKSFLDTLSSLSGFDDQDLFGIWSKVYDPLFCESVH 218
            | |:        |||    ||: |    |:  :|:  | : |: |   ||:    :
Db     221 NTTEAAHTSWRNVFLKPIANRLNKY---FDSGYNLTVSDVRSLYYIC--VYEIALRD--- 272

Qy     219 NFTLPSWATEDAMIKLKELSELSLLSLYGIHKQKEKSRLQGGVLVNEILKNMKLATQPQK 278
           |    |  |    :  :  |:|     |      | :   ||  || :  |::
Db     273 NSDFCSLFTPSEFLNFEYDSDLDY--AYWGGPASEWASTLGGAYVNNLANNLRKGVNNAS 330

Qy     279 YKKLVMYSAHDTTVSGLQMALDVYNGVLP--------------------PYASCHMMEL 317
             :|: :   ||: :  :: ||  :  : |                      |:|   : ||
Db     331 DRKVFLAFTHDSQIIPVEAALGFFPDITPEHPLPTDKNIFTYSLKTSSFVPFAGNLITEL 390

Qy     318 YHDKGGHFVEMYYRNETQNEPYPLT 342
           :     :   | |:    : ||||
Db     391 FLCSDNKY---YVRHLVNQQVYPLT 412


Search completed: June  4, 2002, 10:34:36
```

Job time: 232 sec

THIS PAGE BLANK (USPTO)

File ✓

```
FT   CHAIN        32    381       PROSTATIC ACID PHOSPHATASE.
FT   DISULFID    160    371
FT   DISULFID    346    350
FT   ACT_SITE     43     43
FT   ACT_SITE     85     85       BY SIMILARITY.
FT   CARBOHYD     93     93       N-LINKED (GLCNAC. . .).
FT   CARBOHYD    219    219       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    332    332       N-LINKED (GLCNAC. . .).
SQ   SEQUENCE   381 AA;  43850 MW;  5EEBFF67B062FF76 CRC64;


  Query Match             88.3%;  Score 1818;  DB 1;  Length 381;
  Best Local Similarity   88.5%;  Pred. No. 3.7e-137;
  Matches  337;  Conservative   24;  Mismatches   20;  Indels    0;  Gaps    0;

Qy       1 MGAVPLPLSPTASLSLGFLLLLSLCLDPGQAKELKFVTLVFRHGDRGPIETFPTDPITES 60
           | ||||  |  ||||:||||||||| ||||||||||||||||||||||||||| ||| ||
Db       1 MRAVPLHLVGTASLTLGFLLLLSLRLDPGQAKELKFVTLVFRHGDRGPIETFPNDPIKES 60

Qy      61 SWPQGFGQLTQWGMEQHYELGSYIRKRYGRFLNDTYKHDQIYIRSTDVDRTLMSAMTNLA 120
           ||||||||||:||| ||||||||||:|||||||::|||||:|||||||||||||||||||
Db      61 SWPQGFGQLTKWGMGQHYELGSYIRRRYGRFLNNSYKHDQVYIRSTDVDRTLMSAMTNLA 120

Qy     121 ALFPPEGISIWNPRLLWQPIPVHTVSLSEDRLLYLPFRDCPRFEELKSETLESEEFLKRL 180
           ||||||||||||||||||||||||||||||||||||||||||||:|||||||:|||||||
Db     121 ALFPPEGISIWNPRLLWQPIPVHTVSLSEDRLLYLPFRDCPRFQELKSETLKSEEFLKRL 180

Qy     181 HPYKSFLDTLSSLSGFDDQDLFGIWSKVYDPLFCESVHNFTLPSWATEDAMIKLKELSEL 240
           |||||:||| |||||:||||| |||::||||:|||||||  :||||||| ||||||||
Db     181 QPYKSFIDTLPSLSGFEDQDLFEIWSRLYDPLYCESVHNFTFRTWATEDAMTKLKELSEL 240

Qy     241 SLLSLYGIHKQKEKSRLQGGVLVNEILKNMKLATQPQKYKKLVMYSAHDTTVSGLQMALD 300
           |||||||||||||||||||||||||||||||||||| :||:||||:||||||||||:
Db     241 SLLSLYGIHKQKEKSRLQGGVLVNEILKNMKLATQPQKARKLIMYSAYDTTVSGLQMALE 300

Qy     301 VYNGVLPPYASCHMMELYHDKGGHFVEMYYRNETQNEPYPLTLPGCTHSCPLEKFAELLD 360
           .  :|||:|||||||||:|||| | || ||||||||||||||||||||||||||||||||
Db     301 LYNGLLPPYASCHIMELYQDNGGTFVEMYYRNETQNEPYPLTLPGCTHSCPLEKFAELLD 360

Qy     361 PVIPQDWATECMATSSHQGTV 381
           |||||||||||| ||:|| ::
Db     361 PVIPQDWATECMGTSNHQASL 381


RESULT   2
PPAP_HUMAN
ID   PPAP_HUMAN     STANDARD;      PRT;   386 AA.
AC   P15309;
DT   01-APR-1990 (Rel. 14, Created)
DT   01-MAR-1992 (Rel. 21, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Prostatic acid phosphatase precursor (EC 3.1.3.2).
GN   ACPP.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=92272747; PubMed=1375464;
RA   Sharief F.S., Li S.S.-L.;
RT   "Structure of human prostatic acid phosphatase gene.";
RL   Biochem. Biophys. Res. Commun. 184:1468-1476(1992).
RN   [2]
RP   SEQUENCE FROM N.A., PARTIAL SEQUENCE, DISULFIDE BONDS, AND ACTIVE
RP   SITE.
RX   MEDLINE=91115848; PubMed=1989985;
RA   van Etten R.L., Davidson R., Stevis P.E., Macarthur H., Moore D.L.;
RT   "Covalent structure, disulfide bonding, and identification of
RT   reactive surface and active site residues of human prostatic acid
RT   phosphatase.";
RL   J. Biol. Chem. 266:2313-2319(1991).
RP   SEQUENCE FROM N.A.
RX   MEDLINE=89228054; PubMed=2712834;
RA   Sharief F.S., Lee H., Leuderman M.M., Lundwall A., Deaven L.L.,
RA   Lee C.-L., Li S.S.-L.;
RT   "Human prostatic acid phosphatase: cDNA cloning, gene mapping and
RT   protein sequence homology with lysosomal acid phosphatase.";
RL   Biochem. Biophys. Res. Commun. 160:79-86(1989).
RN   [4]
RP   SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC   TISSUE=Prostate;
RX   MEDLINE=88312981; PubMed=2842184;
RA   Vihko P., Virkkunen P., Henttu P., Roiko K., Solin T., Huhtala M.L.;
RT   "Molecular cloning and sequence analysis of cDNA encoding human
RT   prostatic acid phosphatase.";
RL   FEBS Lett. 236:275-281(1988).
RN   [5]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Prostate;
RX   MEDLINE=90370491; PubMed=2395659;
RA   Tailor P.G., Govindan M.V., Patel P.C.;
RT   "Nucleotide sequence of human prostatic acid phosphatase determined
RT   from a full-length cDNA clone.";
RL   Nucleic Acids Res. 18:4928-4928(1990).
RN   [6]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=95038536; PubMed=7951074;
RA   Sharief F.S., Li S.S.-L.;
RT   "Nucleotide sequence of human prostatic acid phosphatase ACPP gene,
RT   including seven Alu repeats.";
RL   Biochem. Mol. Biol. Int. 33:561-565(1994).
RN   [7]
RP   X-RAY CRYSTALLOGRAPHY (2.9 ANGSTROMS).
RX   MEDLINE=99023966; PubMed=9804805;
RA   Lacount M.W., Handy G., Lebioda L.;
RT   "Structural origins of L(+)-tartrate inhibition of human prostatic
RT   acid phosphatase.";
RL   J. Biol. Chem. 273:30406-30409(1998).
CC   -!- CATALYTIC ACTIVITY: An orthophosphoric monoester + H(2)O = an
CC       alcohol + phosphate.
CC   -!- SUBUNIT: HOMODIMER.
CC   -!- SIMILARITY: BELONGS TO THE HISTIDINE ACID PHOSPHATASE FAMILY.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; M97589; AAA60021.1; -.
DR   EMBL; M97580; AAA60021.1; JOINED.
DR   EMBL; M97581; AAA60021.1; JOINED.
DR   EMBL; M97582; AAA60021.1; JOINED.
DR   EMBL; M97583; AAA60021.1; JOINED.
DR   EMBL; M97584; AAA60021.1; JOINED.
DR   EMBL; M97585; AAA60021.1; JOINED.
DR   EMBL; M97586; AAA60021.1; JOINED.
DR   EMBL; M97587; AAA60021.1; JOINED.
DR   EMBL; M97588; AAA60021.1; JOINED.
DR   EMBL; M34840; AAA69694.1; -.
DR   EMBL; M24902; AAA60022.1; -.
DR   EMBL; X52174; CAA36422.1; -.
DR   EMBL; X53605; CAA37673.1; -.
DR   EMBL; U07097; AAB60640.1; -.
DR   EMBL; U07083; AAB60640.1; JOINED.
DR   EMBL; U07085; AAB60640.1; JOINED.
DR   EMBL; U07086; AAB60640.1; JOINED.
DR   EMBL; U07088; AAB60640.1; JOINED.
DR   EMBL; U07091; AAB60640.1; JOINED.
DR   EMBL; U07092; AAB60640.1; JOINED.
DR   EMBL; U07093; AAB60640.1; JOINED.
DR   EMBL; U07095; AAB60640.1; JOINED.
```

THIS PAGE BLANK (USPTO)

```
DR   PIR; A32419; A32419.
DR   PIR; S01331; S01331.
DR   PIR; S11147; S11147.
DR   PIR; JH0610; JH0610.
DR   PDB; 2HPA; 16-SEP-98.
DR   MIM; 171790; -.
DR   InterPro; IPR000560; His_acid_phosphtse.
DR   Pfam; PF00328; acid_phosphat; 1.
DR   PROSITE; PS00616; HIS_ACID_PHOSPHAT_1; 1.
DR   PROSITE; PS00778; HIS_ACID_PHOSPHAT_2; 1.
KW   Hydrolase; Glycoprotein; Signal; 3D-structure.
FT   SIGNAL        1     32
FT   CHAIN        33    386       PROSTATIC ACID PHOSPHATASE.
FT   DISULFID    161    372
FT   DISULFID    215    313
FT   DISULFID    347    351
FT   ACT_SITE     44     44       NUCLEOPHILIC ACCEPTOR (BY SIMILARITY).
FT   ACT_SITE     86     86       BY SIMILARITY.
FT   ACT_SITE    289    289       PROTON DONOR.
FT   CARBOHYD     94     94       N-LINKED (GLCNAC. . .).
FT   CARBOHYD    220    220       N-LINKED (GLCNAC. . .).
FT   CARBOHYD    333    333       N-LINKED (GLCNAC. . .).
FT   CONFLICT     15     24       SLGFLFLLFF -> ALASCFCFFC (IN REF. 3
FT                                AND 4).
FT   CONFLICT     15     24       SLGFLFLLFF -> AFASCFCFFC (IN REF. 5).
FT   CONFLICT     46     46       D -> H (IN REF. 5).
FT   CONFLICT     66     73       GFGQLTQL -> WIWPTHPA (IN REF. 4).
FT   CONFLICT     66     73       GFGQLTQL -> RIWPTHPA (IN REF. 4).
FT   CONFLICT     95     95       E -> D (IN REF. 3).
FT   CONFLICT    116    116       A -> R (IN REF. 3).
FT   CONFLICT    139    139       Q -> E (IN REF. 5).
FT   CONFLICT    157    157       P -> R (IN REF. 5).
FT   CONFLICT    212    212       P -> A (IN REF. 4).
FT   CONFLICT    215    215       C -> S (IN REF. 3).
FT   CONFLICT    294    294       S -> T (IN REF. 3).
FT   CONFLICT    372    372       C -> V (IN REF. 3).
FT   CONFLICT    383    383       D -> N (IN REF. 5).
SQ   SEQUENCE   386 AA;  44566 MW;  EF81E11DFAECADEA CRC64;
```

```
  Query Match             80.7%;  Score 1661.5;  DB 1;  Length 386;
  Best Local Similarity   81.6%;  Pred. No. 1e-124;
  Matches  311;  Conservative   28;  Mismatches   41;  Indels    1;  Gaps    1;

Qy       1 MGAVPLPLSPTASLSLGFLLLLSLCLDPG-QAKELKFVTLVFRHGDRGPIETFPTDPITE 59
           | | || |:  |||||||| ||   ||    |||||||||||||||| ||:||||||| |
Db       1 MRAAPLLLARAASLSLGFLFLLFFWLDRSVLAKELKFVTLVFRHGDRSPIDTFPTDPIKE 60

Qy      60 SSWPQGFGQLTQWGMEQHYELGSYIRKRYGRFLNDTYKHDQIYIRSTDVDRTLMSAMTNL 119
           |||||||||||| |||||||| |||||| :|||::|||:|:|||||||||||||||||
Db      61 SSWPQGFGQLTQLGMEQHYELGEYIRKRYRKFLNESYKHEQVYIRSTDVDRTLMSAMTNL 120

Qy     120 AALFPPEGISIWNPRLLWQPIPVHTVSLSEDRLLYLPFRDCPRFEELKSETLESEEFLKR 179
           ||||||||:||||| |||||||||||| |||||:|||||||:||||:||:||||:|||| ||
Db     121 AALFPPEGVSIWNPILLWQPIPVHTVPLSEDQLLYLPFRNCPRFQELESETLKSEEFQKR 180

Qy     180 LHPYKSFLDTLSSLSGFDDQDLFGIWSKVYDPLFCESVHNFTLPSWATEDAMIKLKELSE 239
           ||||| |: ||  |||   |||||||||||||||:|||||||||||||||| | ||:||||
Db     181 LHPYKDFIATLGKLSGLHGQDLFGIWSKVYDPLYCESVHNFTLPSWATEDTMTKLRELSE 240

Qy     240 LSLLSLYGIHKQKEKSRLQGGVLVNEILKNMKLATQPQKYKKLVMYSAHDTTVSGLQMAL 299
           ||||||||||||||||||||||||||| :|| |||   ||||:|||||||||||||||||
Db     241 LSLLSLYGIHKQKEKSRLQGGVLVNEILNHMKRATQIPSYKKLIMYSAHDTTVSGLQMAL 300

Qy     300 DVYNGVLPPYASCHMMELYHDKGGHFVEMYYRNETQNEPYPLTLPGCTHSCPLEKFAELL 359
           |||||:||||||||: ||| :|| :|||||||||||:||||| ||||: |||||:||||:
Db     301 DVYNGLLPPYASCHLTELYFEKGEYFVEMYYRNETQHEPYPLMLPGCSPSCPLERFAELV 360

Qy     360 DPVIPQDWATECMATSSHQGT 380
           |||||||:|||| |:|||||
Db     361 GPVIPQDWSTECMTTNSHQGT 381
```

```
RESULT   3
PPAL_HUMAN
ID   PPAL_HUMAN     STANDARD;      PRT;   423 AA.
AC   P11117;
DT   01-JUL-1989 (Rel. 11, Created)
DT   01-JUL-1989 (Rel. 11, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Lysosomal acid phosphatase precursor (EC 3.1.3.2) (LAP).
GN   ACP2.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC   TISSUE=Placenta;
RX   MEDLINE=89052645; PubMed=3191910;
RA   Pohlmann R., Krentler C., Schmidt B., Schroeder W., Lorkowski G.,
RA   Culley J., Mersmann G., Geier C., Waheed A., Gottschalk S.,
RA   Grzeschik K.H., Hasikik A., von Figura K.;
RT   "Human lysosomal acid phosphatase: cloning, expression and
RT   chromosomal assignment.";
RL   EMBO J. 7:2343-2350(1988).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Leukocyte;
RX   MEDLINE=89377828; PubMed=2776754;
RA   Geier C., von Figura K., Pohlmann R.;
RT   "Structure of the human lysosomal acid phosphatase gene.";
RL   Eur. J. Biochem. 183:611-616(1989).
CC   -!- CATALYTIC ACTIVITY: An orthophosphoric monoester + H(2)O = an
CC       alcohol + phosphate.
CC   -!- SUBCELLULAR LOCATION: Lysosomal.
CC   -!- SIMILARITY: BELONGS TO THE HISTIDINE ACID PHOSPHATASE FAMILY.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; X12548; CAA31064.1; -.
DR   EMBL; X15525; CAA33542.1; -.
DR   EMBL; X15526; CAA33542.1; JOINED.
DR   EMBL; X15527; CAA33542.1; JOINED.
DR   EMBL; X15528; CAA33542.1; JOINED.
DR   EMBL; X15529; CAA33542.1; JOINED.
DR   EMBL; X15530; CAA33542.1; JOINED.
DR   EMBL; X15531; CAA33542.1; JOINED.
DR   EMBL; X15532; CAA33542.1; JOINED.
DR   EMBL; X15533; CAA33542.1; JOINED.
DR   EMBL; X15534; CAA33542.1; JOINED.
DR   EMBL; X15535; CAA33542.1; JOINED.
DR   PIR; S06167; S06167.
DR   HSSP; P20646; 1RPA.
DR   MIM; 171650; -.
DR   InterPro; IPR000560; His_acid_phosphtse.
DR   Pfam; PF00328; acid_phosphat; 1.
DR   PROSITE; PS00616; HIS_ACID_PHOSPHAT_1; 1.
DR   PROSITE; PS00778; HIS_ACID_PHOSPHAT_2; 1.
KW   Hydrolase; Signal; Glycoprotein; Lysosome.
FT   SIGNAL        1     30
FT   CHAIN        31    423       LYSOSOMAL ACID PHOSPHATASE.
FT   DISULFID    159    370       BY SIMILARITY.
FT   DISULFID    212    310       BY SIMILARITY.
FT   DISULFID    345    349       BY SIMILARITY.
FT   ACT_SITE     42     42       NUCLEOPHILIC ACCEPTOR (BY SIMILARITY).
FT   ACT_SITE     84     84       BY SIMILARITY.
FT   ACT_SITE    286    286       PROTON DONOR (BY SIMILARITY).
FT   CARBOHYD     92     92       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    133    133       N-LINKED (GLCNAC. . .) (POTENTIAL).
```

THIS PAGE BLANK (USPTO)